The Dissertation Committee for Javier Alejandro Morales Delgado
certifies that this is the approved version of the following dissertation:

Least action principles with applications to gradient flows and kinetic equations.

Committee:

Luis A. Caffarelli, Supervisor

Alessio Figalli

Irene Gamba

Francesco Maggi

Alexis F. Vasseur

Least action principles with applications to gradient flows and kinetic equations.

by

Javier Alejandro Morales Delgado, B.A, M.A.

DISSERTATION

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

DOCTOR OF PHILOSOPHY

THE UNIVERSITY OF TEXAS AT AUSTIN

May 2017

Dedicated to my Ph.D. advisor, Alessio Figalli.

Least action principles with applications to gradient flows and kinetic equations.

Publication No. ______________

Javier Alejandro Morales Delgado, Ph.D.
The University of Texas at Austin, 2017

Supervisor: Luis A. Caffarelli

This thesis introduces a variational formulation for a family of kinetic reaction-diffusion and their connection to Lagrangian dynamical systems. Such a formulation uses a new class of transportation costs between positive measures, and it generalizes the notion of gradient flows. We use this class to build solutions to reaction-diffusion equations with drift subject to general Dirichlet boundary condition via an extension of De Giorgi's interpolation method for the entropy functional. In 2010, Alessio Figalli and Nicola Gigli introduced a transportation cost that can be used to obtain parabolic equations with drift subject to Dirichlet boundary condition. However, the drift and the boundary condition are coupled in their work. The costs we introduce allow the drift and the boundary condition to be decoupled.

Additionally, we use this variational formulation to obtain well-posedness, stability, and convergence to equilibrium for the homogeneous Vicsek model and to show the emergence of phase concentration for the Kuramoto Sakaguchi equation subject to a strong coupling force. Provided this coupling force is sufficiently large, we show that there exists a time-dependent interval such that the oscillator's probability density converges to zero uniformly in its complement. The length of this interval is quantified as a function of the coupling force and the diameter of the support of the natural frequency distribution. By doing this, we show that the diameter of the interval can be made arbitrarily small by choosing the force sufficiently large.

Table of Contents

Chapter 1

Introduction

In this thesis, we establish a connection between the theory of optimal transportation and Hamilton's analogy between geometrical optics and classical mechanics. Such a connection allows us to develop a heuristic framework for a large family of kinetic and reaction-diffusion equations with boundary conditions. We use this framework as an intuitive guideline to obtain well-posedness, stability, and dynamics for some of these equations. The starting point is the theory of gradient flows in the space of probability measures. This theory has recently been studied extensively as an application of the optimal transport problem.

The study of the optimal transport problem was initiated by Gaspard Monge in 1781 with his work *"Mémoire sur la théorie des déblais et des remblais"*. The problem considered therein is to determine the assignment between the locations where materials should be extracted and the final positions to which they should be transported. This assignment should minimize the total transportation cost. For each unit of mass, such a cost is given by the distance traveled. Monge discovered that the trajectories followed by the masses should go along straight lines that are orthogonal to a family of surfaces.

To state Monge's problem using mordern terminology we introduce the following notation: given complete and separable metric spaces (X, d) and (Y, d) we denote by $\mathbb{P}(X)$ and $\mathbb{P}(Y)$ the set of probability measures on X and Y. If $T : X \to Y$ is a Borel map, and $\mu \in \mathbb{P}(X)$ is a probability measure, we will say that the measure $T_{\#}\mu \in \mathbb{P}(Y)$ defined by

$$T_{\#}\mu(B) = \mu(T^{-1}(B)) \quad \forall B \subset Y \quad \text{Borel},$$

is the push forward of μ through T.

Let $c : X \times Y \to \mathbb{R} \times \{\infty\}$ be a cost function. The Monge version of the optimal transport problem for the cost c is the following: Given μ and ν be in $\mathbb{P}(X)$ and $\mathbb{P}(Y)$, minimize

$$T \to \int_X c(x, T(x))\, d\mu, \tag{1.0.1}$$

among all transport maps T satisfying $T_{\#}\mu = \nu$.

The work of Monge has been revisited by many researchers. Among them was Leonid Vitaliyevich Kantorovich in 1938. In 1975 Kantorovich was awarded the Nobel Prize in Economics, jointly with Tjalling Koopmans, "for their contributions to the theory of optimum allocation of resources."

In his work Kantorovich obtained a duality theorem for the optimal transportation problem that plays a crucial role in the mordern theory. Such a dual formulation produces a potential whose level sets are the surfaces discovered by Monge. After the establishment of the connection between the works of Monge and Kantorovich, the problem received the name of the Monge-Kantorovich problem.

The Kantorovich formulation of the optimal transport problem is the following: Given μ in $\mathbb{P}(X)$ and ν in $\mathbb{P}(Y)$, minimize

$$\gamma \to \int_{X\times Y} c(x,y)\, d\gamma,$$

in the set $ADM(\mu,\nu)$ of all transport plans γ in $\mathbb{P}(X,Y)$ satisfying $\pi^1_{\#}\gamma = \mu$ and $\pi^2_{\#}\gamma = \nu$. Here, π^1 and π^2 are the natural projections from $X \times Y$ onto X and Y respectively.

The dual formulation discovered by Kantorovich can be stated as follows: Given μ and ν, maximize the value of

$$\int \varphi\, d\mu + \int \psi\, d\nu,$$

among all functions $\varphi \in L^1(\mu)$ and $\psi \in L^1(\nu)$ such that

$$\varphi(x) + \psi(y) \leq c(x,y), \quad \forall x \in X \text{ and } y \in Y.$$

These two problems are related by the fact that

$$\inf_{\gamma \in Adm(\mu,\nu)} \int c(x,y)\, d\gamma = \sup_{\varphi \in L^1(\mu), \psi \in L^1(\nu)} \int \varphi\, d\mu + \int \psi\, d\nu.$$

Yann Brenier [12, 13] provided a solution to the Monge problem in $\mathbb{R}^d$ when $c(x,y) = |x-y|^2$. He proved that whenever μ is an absolutely continuous measure in $\mathbb{P}(\mathbb{R}^d)$ satisfying $\int |x|^2 d\mu < \infty$ and ν is a measure in $\mathbb{P}(\mathbb{R}^d)$ satisfying $\int |x|^2 d\nu < \infty$, there exists a unique map T minimizing (1.0.1). Moreover, he showed that T can be recovered by taking the gradient of a convex function.

For the quadratic cost the optimal transport problem induces a distance in the space of probability measures with finite second moments. We will denote this space by $\mathbb{P}_2(\mathbb{R}^d)$. This distance is called the Wasserstein distance. The Wasserstein distance between two measures μ and ν in $\mathbb{P}_2(\mathbb{R}^d)$ is given by

$$W_2(\mu,\nu) := \left(\inf_\gamma \int |x-y|^2 d\gamma \right)^{\frac{1}{2}}.$$

Here, γ ranges in the set of all transport plans in $\mathbb{R}^d \times \mathbb{R}^d$ with marginals μ and ν. This distance turns the space of all probability measures in $\mathbb{P}(\mathbb{R}^d)$ with finite second moments into a complete metric space. We will denote such space by $(\mathbb{P}_2(\mathbb{R}^d), W_2)$.

In the quadratic cost case, Benamou and Brenier [11] introduced a hydrodynamic formulation of the optimal transport problem: Given μ and ν in $\mathbb{P}_2(\mathbb{R}^d)$, minimize

$$\int_0^1 \int |v_t|^2 \, d\mu_t \, dt,$$

among all vector fields v_t and measures μ_t indexed in the interval $[0,1]$ that satisfy the constraint:

$$\begin{cases} \frac{d}{dt}\int \zeta \, d\mu_t = \int \nabla\zeta v_t \, d\mu_t \quad \forall \zeta \in C^\infty(\mathbb{R}) \text{ and } \forall t \in [0,1], \\ \mu_0 = \mu, \\ \mu_1 = \nu. \end{cases} \tag{1.0.2}$$

The connection with the optimal transport problem is given by the Benamou-Brenier formula:

$$W_2(\mu,\nu) := \sqrt{\min_{v_t,\mu_t} \int_0^1 \int |v_t|^2 \, d\mu_t dt}.$$

Here the minimum is taken among all vector fields v_t and measures μ_t index in $[0,1]$ that satisfy (1.0.2).

The result of Brenier in [12, 13] established a link between optimal transportation and fluid mechanics and the theory of Monge-Ampère equations. This attracted the attention of the community working on partial differential equations. Many researchers from this community have worked on improving our understanding of the optimal transportation problem and the scope of its applications.

The application that is most relevant to this thesis is the one discovered by Felix Otto. He achieved an important step in which he introduced a formalism in which the space of probability measures under the Wasserstein distance can be seen as a Riemannian manifold.

This formalism provides a powerful heuristic framework that allows using the theory of optimal transportation to study evolution equations. More precisely, one of the most remarkable achievements of [66, 83] has been to illustrate that many evolution equations are gradient flows of entropy functionals. By doing formal computations in the infinite-dimensional Riemannian structure of $(\mathbb{P}_2(\mathbb{R}^d), W_2)$, Otto was able to obtain quantitative convergence rates and stability for solutions of these gradient flows. In general:

> *"The strategy of applying geometric methods to the infinite-dimensional problems is as follows. Having established certain facts in the finite-dimensional situation, one uses the results to formulate the corresponding facts for the infinite-dimensional case... These final results often can be proved directly, leaving aside the difficult questions of foundations for the intermediate steps (V. I. Arnold and B. A. Khesin, Topological Methods in Hydrodynamics [7])."*

This approach has successfully been used by Otto and Villani to study functional inequalities [84]. The results of Otto [83] rely strongly on having uniform convexity bounds, and nowadays they follow by the established theory [6].

The objective of this thesis is to extend these applications to situations where the energy functional may not have a universal convexity lower bound, the model may not preserve mass, or a gradient flow structure may not be available. Such situations include reaction-diffusion equations subject to boundary conditions, nonlocal diffusion equations, kinetic equations and models of collective dynamics.

We extend these applications to kinetic and reaction-diffusion equations with boundary conditions where gradient flow structures are not available. We do so by introducing a new formalism. This formalism is the result of combining generalizations of the ideas of Felix Otto with the analogy between geometrical optics and mechanics introduced by Hamilton at the level of action functionals in the space of paths of positive measures. The essential tool that allows us to do formal computation in such space is the formal differential structure in the space of paths introduced by Milnor in [78, Part III].

The Lagrangian formulation of classical mechanics was introduced Joseph Louis Lagrange in 1788. In Lagrangian mechanics, the trajectory of a particle system is a local minimizer of an action functional in the space of paths. This formulation allows the treatment of systems subjected to holonomic constraints. There, the system is confined to a submanifold in the configuration space. This formulation allows one to find equations of motions that are independent of the choice of coordinates in such manifolds.

Hamilton introduced the principle of least action in 1827 in his treatise Theory of Systems of Rays, in which he developed the Hamiltonian formulation of mechanics. This work brought together classical mechanics and optics and helped to establish the wave theory of light. Just like Monge, he discovered that for isotropic mediums, rays move in the direction orthogonal to the front of light waves. The work of Hamilton has been extended by many scientists including Liouville, Jacobi, Darboux, Poincare, Kolmogorov, and Arnold. These extensions were the starting point of new mathematical fields such as Dynamical systems and Symplectic Geometry.

A particular important extension is the one developed by John Mather. In his work [75, 76], he considers the problem of not only studying action-minimizing curves but also studying action minimizing measures in phase space. Under some conditions on the action, Mather proved a result in which certain action-minimizing measures are concentrated in Lipschitz graphs. This is a major step in the direction of understanding the long time behavior of Lagrangian dynamical systems and the existence of periodic orbits.

One of the objectives of this thesis is to begin a program in which the tools from Lagrangian mechanical systems can be used to study the dynamics of some reaction-diffusion and kinetic equations.

In [38] Albert Fathi made important connections between the theory of viscosity solutions of the Hamilton-Jacobi equations and the work of Mather. The viscosity solutions of the stationary Hamilton-Jacobi equations are strongly related to the Lipschitz graphs in which the action minimizing measures of

Mather are concentrated. Nowadays, the study of this connection is known as weak KAM theory. In particular, Fathi showed that viscosity solutions of the stationary Hamilton-Jacobi equation are calibrated variationally with a system of action minimizing curves. In the heuristic framework this thesis introduces, solutions to some kinetic and reaction-diffusion equations are the calibrated curves corresponding to a functional in the space of positive measures that solves a suitable Hamilton-Jacobi equation.

The rest of the introduction is devoted to the exposition of the gradient flows in the formal Riemannian structure introduced by Felix Otto, its generalization based on weak KAM structures for action functionals in the space of paths of measures and its applications to reaction-diffusion and kinetic equations. In Section 1 we explain the relationship between gradient flows and weak KAM structures on finite dimension and how it can be used to obtain information about the dynamics. In Section 2 we introduce the formal Riemannian structure of Felix Otto. In Section 3 we discuss the formal differential structure of the space of paths introduced by John Milnor. In Section 4 we explain how the homogeneous Vicsek model can be seen as a gradient flow in this formal Riemannian structure and how this interpretation can be used to obtain stability estimates and convergence rates. In Section 5 we introduce the formal weak KAM structure of action functionals in the space of paths of probability measures. We also explain how this structure was used to obtain information about the long-term behavior of the Kuramoto Sakaguchi equation and how other models such as the Stochastic Vicsek model and the Cucker-Smale model can be interpreted according to this structure. In Section 6 we introduce a formal weak KAM structure for reaction-diffusion equations with boundary conditions and explain how it can be used to get the existence of solutions via the De Giorgi minimizing movement scheme. These sections are written in an expository language, and most of the statements are just formal. The purpose is to give the intuition that was used to derive the results in the subsequent chapters.

These chapters are based on papers [44, 53, 67, 79] that show rigorous results about several evolution equations. In Chapter 2 we prove rigorously that one can use the De Giorgi minimizing movement scheme of the relative entropy with a new family of transportation costs to produce solutions to a large class of reaction-diffusion equations subject to boundary conditions. In Chapter 3 we prove well-posedness, stability, and convergence to equilibrium for the homogeneous Vicsek model for $d \geq 3$. In Chapter 4 we prove convergence with a quantitative rate for the homogeneous Vicsek model in the plane. Finally, in Chapter 5 we prove the emergence of a concentration phenomena for the Kuramoto-Sakaguchi equation in a large coupling strength regime.

1.1 The finite dimensional case

In this section, we discuss the theory of gradient flows in $\mathbb{R}^d$ and how the theory of geometrical optics can be used to extend it. After stating the gradient flow problem, we explain how one can use an implicit Euler method based on the minimizing movement scheme to get existence of the gradient flow problem with prescribed initial data. Then, we analyze how under the strict convexity assumptions on the energy one can obtain quantitative convergence rates and contraction of solutions.

Afterwards, we introduce geometrical optics and we explain how it generalized the concept of gradient flows. Finally, we discuss how the theory of Aubry, Fathi, Mañe, and Mather can be used to study the dynamics of the systems of rays. The final goal is to apply these ideas to kinetic and reaction-diffusion equations in the subsequent chapters.

Given a function $E : \mathbb{R}^d \to \mathbb{R}$ in $C^2(\mathbb{R}^d)$ and a point x_0 in $\mathbb{R}^d$, the gradient flow of E starting at x_0 is given by a curve $x : [0, \infty) \to \mathbb{R}^d$ satisfying

$$x(0) = x_0,$$

and

$$\dot{x} = -\nabla_x E \quad \forall t \in [0, \infty).$$

1.1.1 The minimizing movement scheme

One way to build approximate solutions to the gradient flow of E starting at x_0 is by following an implicit Euler scheme. Given a time step $\tau > 0$, this scheme produces a sequence of points $\{x_{k\tau}\}_{k=0}^{\infty}$ where $x_{k\tau}$ is inductively determined from $x_{(k-1)\tau}$ by setting

$$\frac{x_{k\tau} - x_{(k-1)\tau}}{\tau} = -\nabla_{x_{k\tau}} E. \tag{1.1.3}$$

The scheme is called implicit because the gradient of E is evaluated at $x_{k\tau}$ instead of $x_{(k-1)\tau}$. This choice allows us to restate the above condition as a minimization problem. Indeed, (1.1.3) can be rewritten as:

$$\nabla_{x=x_{k\tau}} \left(\frac{|x - x_{(k-1)\tau}|^2}{2\tau} + E \right) = 0.$$

If $D^2 E \geq \lambda \,\mathrm{Id}$ for some λ in $\mathbb{R}$, for sufficiently small τ, the above condition is equivalent to the fact $x_{k\tau}$ minimizes

$$x \to \frac{|x - x_{(k-1)\tau}|^2}{2\tau} + E(x).$$

The advantage of restating (1.1.3) as a minimization problem is that it allows us to obtain energy estimates at the level of discrete solutions and show compactness of the family of functions $\{x^\tau(t)\}_{\tau \in (0,1)}$. These functions are defined by setting $x^\tau(t) := x_{(k-1)\tau}$, whenever t is in $[(k-1)\tau, k\tau)$. Indeed, by minimality of $x_{k\tau}$ we have

$$\frac{|x_{k\tau} - x_{(k-1)\tau}|^2}{2\tau} + E(x_{k\tau}) \leq \frac{|x_{(k-1)\tau} - x_{(k-1)\tau}|^2}{2\tau} + E(x_{(k-1)\tau}) = E(x_{(k-1)\tau}).$$

Thus, summing the above relationship from $k = n$ to $k = m$ yields,

$$\sum_{k=n}^{m} \frac{|x_{k\tau} - x_{(k-1)\tau}|^2}{2\tau} \leq E(x_{m\tau}) - E(x_{(n-1)\tau}).$$

Hence, by Jensen inequality,

$$\frac{1}{2\tau(n-m)} \left| \sum_{k=n}^{m} x_{k\tau} - x_{(k-1)\tau} \right|^2 \leq E(x_{m\tau}) - E(x_{(n-1)\tau}).$$

Consequently, for any natural numbers $m \geq n \geq 1$, we have

$$|x_{m\tau} - x_{(n-1)\tau}| \leq \sqrt{2(n-m)\tau[E(x_{m\tau}) - E(x_{(n-1)\tau})]},$$

which easily yields equicontinuity of the family $\{x^\tau(t)\}_{\tau\in(0,1)}$. This allows us to obtain a limiting curve as $\tau \to 0$ via the Ascoli-Arzelà Theorem.

The analog of this method was used for the first time in $(\mathbb{P}_2(\mathbb{R}^d), W_2)$ by Jordan, Kinderlerher, and Otto [66] to build solutions to the Fokker-Planck equation. Nowadays, it is well understood that by using different energy functionals in $(\mathbb{P}_2(\mathbb{R}^d), W_2)$ one can use the method to build solution to equations of the form

$$\partial_t \rho(t) = \operatorname{div}(\nabla\rho - \rho V - \rho(\nabla W * \rho)).$$

We refer the reader to [5], for an excellent introduction in this topic.

1.1.2 Dynamics for strictly convex gradient flows

Here, we show that if

$$D^2 E \geq \lambda \,\mathrm{Id} \quad \text{for some} \quad \lambda > 0, \tag{1.1.4}$$

then solutions to the gradient flow converge exponentially fasts to the minimum of E. That is, we have

$$|x_t - x_{\min}| \leq \sqrt{\frac{2}{\lambda} E(x_0)} e^{-\lambda t}.$$

By (1.1.4) and the assumption that E is in $C^2(\mathbb{R}^d)$, we have that there exists a unique $x_{\min}$ in $\mathbb{R}^d$ satisfying

$$\begin{cases} E(x) \geq E(x_{\min}) & \forall x \in \mathbb{R}^d, \\ \nabla_{x_{\min}} E = 0, \\ \lim_{k\to\infty} |\nabla_{y_k} E| = 0 & \implies y_k \to x_{\min}. \end{cases} \tag{1.1.5}$$

Without loss of generality, let us assume

$$E(x_{\min}) = 0.$$

By direct computation,

$$\frac{d}{dt} E(x_t) = \langle \nabla_{x_t} E, \dot{x}_t \rangle = -|\nabla_{x_t} E|^2. \tag{1.1.6}$$

Hence, E is non-increasing along x. Using condition (1.1.13) and Gronwall's inequality, we find

$$\frac{d}{dt} |\nabla_{x_t} E|^2 = -2\langle D^2_{x_t} E \nabla_{x_t} E, \nabla_{x_t} E \rangle \leq -2\lambda |\nabla_{x_t} E|^2.$$

Thus, for any $T \geq 0$.

$$|\nabla_{x_t} E|^2 \leq |\nabla_{x_T} E|^2 e^{-2\lambda(t-T)} \quad \forall t \in [T, \infty). \tag{1.1.7}$$

Integrating, (1.1.6), we obtain

$$\begin{aligned} E(x_0) - E(x_t) &= \int_0^t |\nabla_{x_t} E|^2 ds \\ &\leq |\nabla_{x_0} E|^2 \int_0^t e^{-2\lambda s} ds. \end{aligned}$$

Here, $0 \leq s \leq t$. Thus, letting $t \to \infty$, we get

$$E(x_0) \leq \frac{1}{2\lambda}|\nabla_{x_0} E|^2.$$

Moreover, since x_0 was arbitrary, we conclude

$$E \leq \frac{1}{2\lambda}|\nabla E|^2 \quad \text{in} \quad \mathbb{R}^d. \tag{1.1.8}$$

In particular, combining this with (1.1.6) we obtain

$$\frac{d}{dt}E(x_t) = -|\nabla_{x_t} E|^2 \leq -2\lambda E(x_t).$$

Consequently,

$$E(x_t) \leq e^{-2\lambda t} E(x_t) \quad \forall t \in [0, \infty). \tag{1.1.9}$$

Now, note

$$\begin{aligned} \frac{d}{dt}\frac{1}{2}|x_t - x_{\min}|^2 &= \langle x_t - x_{min}, \dot{x}_t \rangle \\ &= -\langle x_t - x_{min}, \nabla_{x_t} E \rangle \\ &\geq -|x_t - x_{\min}||\nabla_{x_t} E|. \end{aligned} \tag{1.1.10}$$

On the other hand,

$$\frac{d}{dt}\frac{1}{2}|x_t - x_{\min}|^2 = |x_t - x_{\min}|\frac{d}{dt}|x_t - x_{\min}|. \tag{1.1.11}$$

Combining (1.1.10) and (1.1.11), we get

$$\frac{d}{dt}|x_t - x_{\min}| \geq -|\nabla_{x_t} E|.$$

Thus, by (1.1.8),

$$\begin{aligned} \frac{d}{dt}|x_t - x_{\min}| - \sqrt{\frac{2}{\lambda}E(x_t)} &\geq \left(\frac{d}{dt}|x_t - x_{\min}|\right) + \frac{|\nabla_{x_t} E|^2}{\sqrt{2\lambda E(x_t)}} \\ &\geq -|\nabla_{x_t} E| + \frac{|\nabla_{x_t} E|^2}{\sqrt{2\lambda E(x_t)}} \\ &\geq 0. \end{aligned}$$

Hence, the map $t \to |x_t - x_{\min}| - \sqrt{\frac{2}{\lambda}E(x_t)}$ is non-decreasing. Thus, using condition (1.1.5), we see that

$$|x_0 - x_{\min}| - \sqrt{\frac{2}{\lambda}E(x_0)} \leq \limsup_{t\to\infty} |x_t - x_{\min}| - \sqrt{\frac{2}{\lambda}E(x_t)} = 0.$$

Consequently, since x_0 was arbitrary, we conclude

$$|x - x_{\min}| \leq \sqrt{\frac{2}{\lambda}E(x)} \quad \forall x \in \mathbb{R}^d. \tag{1.1.12}$$

Combining this with (1.1.9), yields the convergence to the minimum of E:

$$|x_t - x_{\min}| \leq \sqrt{\frac{2}{\lambda}E(x_0)} e^{-\lambda t}.$$

We refer the reader [84] for more application of the theory of gradient flows to functional inequalities. There Felix Otto and Cedric Villani discuss the ideas of this section and its analog in the space of probability measures over a Riemannian manifold (M, g). In particular, they prove generalization of the Talagrand Inequality and logarithmic Sobolev inequality by interpreting them as the analogue of inequalities (1.1.8) and (1.1.12) for the gradient flow of the relative entropy in $(\mathbb{P}_2(M), W_2)$.

1.1.3 Stability for gradient flows.

Here, we show that solutions to strictly convex gradient flow, that is

$$D^2 E \geq \lambda \,\mathrm{Id} \quad \text{for some} \quad \lambda > 0, \tag{1.1.13}$$

are contractive. More precisely, let x and $\tilde{x}$ be solutions to the gradient flow of E starting at x_0 and $\tilde{x}_0$. Then, we have

$$|x(t) - \tilde{x}(t)| \leq e^{-\lambda t}|x_0 - \tilde{x_0}|. \tag{1.1.14}$$

Indeed, by direct computation

$$
\begin{aligned}
\frac{d}{dt}\frac{1}{2}|x-\tilde{x}|^2 &= \langle x-\tilde{x}, \dot{x}-\dot{\tilde{x}}\rangle \\
&= \langle x-\tilde{x}, \nabla_{\tilde{x}}E\rangle - \langle x-\tilde{x}, \nabla_x E\rangle \\
&= \int_0^1 \frac{d}{ds}\langle x-\tilde{x}, \nabla_{x+s(\tilde{x}-x)}E\rangle ds \\
&= -\int_0^1 \langle x-\tilde{x}, D^2_{(x+s(\tilde{x}-x)}E(x-\tilde{x})\rangle ds \\
&\leq -\lambda|x-\tilde{x}|^2.
\end{aligned}
$$

Hence, the desired result follows.

The analogue of this method was used for the first time in $(\mathbb{P}_2(\mathbb{R}^d), W_2)$ by Felix Otto[83] to show that solutions to a family of Porous medium equations contract exponentially fast in the Wasserstein metric.

1.1.4 Gradient flows and geometrical optics

In this section, we explain the connection between gradient flows and the analogy between geometrical optics and classical mechanics introduced by Hamilton. In Hamilton's theory, light moves in rays that travel from a point x_0 to another point x_{t_0} in the shortest possible time. The way in which lights propagates in a non-isotropic medium is characterized by the indicatrix function L in such a way that if x is a light ray through points x_0 and x_{t_0} at time t_0, then

$$
t_0 = \int_0^{t_0} L(x_s, \dot{x}_s)\, ds = \min_{\alpha} \int_0^{t_0} L(\alpha_s, \dot{\alpha}_s)\, ds, \tag{1.1.15}
$$

where α ranges among all paths satisfying $\alpha(0) = x_0$ and $\alpha(t_0) = x_{t_0}$. Here, $(x, v) \to L(x, v)$ is a symmetric quadratic form in v whose hessian in the v variable has a uniform (in x) positive lower bound, and t_0 is small enough.

Condition (1.1.15) determines the speed at which light rays are allowed to travel. In general mediums, the speed of the light depends on the point and the direction of the ray. If we fix a point x_0 we can look at the points in which the light can travel from x_0 in time less than t_0. The boundary of this set is the wave-front of the point x_0 at time t_0. According to Huygens's principle [8, Huygens's Theorem, chapter 9] the wave front from a point x_0 after time $t+s$ is given by the envelope of the fronts at the time s from all the points in the front of x_0 at time t. Now, define a function $\varphi_{x_0}(x, t)$ by

$$\varphi_{x_0}(x, t) = \inf_{\alpha} \int_0^t L(\alpha_s, \dot{\alpha}_s)\, ds.$$

Here, α ranges among all paths satisfying $\alpha(0) = x_0$ and $\alpha(t) = x$. Then, by definition, the set $\{x \,:\, \varphi_{x_0}(x, t) = t\}$ agrees with the wave-front starting from x_0 at time t. If the medium is homogeneous and isotropic, at any given time, the spatial gradient of φ_{x_0} is perpendicular to the wave-front and co-linear to the rays in which light propagates. For general mediums, the velocity of the ray at x in time t is related to φ_{x_0} through the Legendre transform:

$$\dot{x}(t) = \mathcal{L}_{x_t}(d_{x_t}\varphi_t),$$

where

$$\mathcal{L}_x(p) := \operatorname{argmax}_v \langle p, v\rangle - L(x, v). \tag{1.1.16}$$

(Here, $d_x\varphi_t$ denotes the differential of the function $y \to \varphi_{x_0}(y,t)$ evaluated at point x. We refer the reader to [8] chapter 9, Section A, for a proof of the above relationship. There the indicatrix surface at a given point x is the 1-level set of the map $v \to L(x,v)$. Additionally, the function S_{x_0}, defined in [8] chapter 9, satisfies $\{x \,:\, S_{x_0}(x) = t\} = \{x \,:\, \varphi_{x_0}(x,t) = t\}$.)

In the case where $L(x,v) = \frac{1}{2}|v|^2$, we have that

$$\mathcal{L}_x(d_x\varphi_t) = \nabla_x\varphi_t.$$

Additionally, the function $\varphi := \varphi_{x_0}$ satisfies the equation

$$\partial_t\varphi + H(x, \nabla\varphi) = 0,$$

where

$$H(x,p) = \sup_v \langle p, v\rangle - L(x,v).$$

Let us impose periodic boundary conditions by letting the rays move in the torus $\mathbb{T}^d$ or more generally in a compact Riemannian manifold (M,g). Additionally, suppose H satisfies conditions (i) to (iii) below:

(i) H is in $C^k(M)$.

(ii) For every $K \geq 0$, there is a finite constant $C^*(K)$ such that,

$$H(x,p) \geq K||p||_x + C^*(K).$$

(iii) For every $(x,p) \in T^*M$, the second derivative along the fibers $\frac{\partial^2 H}{\partial p^2}(x,p)$ is positive definite.

Then, as $t \to \infty$ we have that $\partial_t \varphi \to c[H]$, for some constant. This constant is called the critical value of the Hamiltonian H (see [39] for a proof of this convergence).

Hence, $\lim_{t\to} \varphi(x,t) = u(x) - ct$. The function u is called a KAM function and it is calibrated with a system of rays. That is, for each x_0 there exists a curve $x : (-\infty, 0] \to \mathbb{R}$ that is an extremal of the action and satisfies

$$u(x(b)) - u(x(a)) = \int_a^b L(x, \dot{x}) ds + c[H](b-a) \quad \forall a \leq b \leq 0,$$

and

$$x(0) = x_0.$$

Such curves are C^2 and their velocity are given by the function u via the Legendre transform:

$$\dot{x}(t) = \text{argmax}_v \langle d_{x_t} u, v \rangle - L(x_t, v) = \mathcal{L}_x(d_{x_t} u) \; \forall t \in (-\infty, 0].$$

(See [88] for a concise introduction to this topic.)

When $H(x,p) = \frac{|p|^2}{2}$, this relationship translates to

$$\dot{x}(t) = \nabla_{x_t} u,$$

which is in fact a gradient flow.

Conversely, given a differentiable function E and a Lagrangian L, consider the Lagrangian

$$\tilde{L}(x,v) := L(x,v) + H(x, d_x E),$$

and its respective Hamiltonian

$$\begin{aligned}\tilde{H}(x,p) &= \sup_v vp - L(x,v) - H(x, d_x E) \\ &= H(x,p) - H(x, d_x E).\end{aligned}$$

Then, E solves the equation

$$\tilde{H}(x, d_x E) = 0.$$

Hence, E is a so-called KAM function, the critical value of $\tilde{H}$ is 0, and the curves

$$\dot{x}(t) = \mathcal{L}_{x_t}(d_{x_t} E),$$

are calibrated with respect to E.

In Section 5 and 6 we present infinite dimensional analogues of the relationship between KAM functions and its calibrated systems of rays. We will call this a weak KAM structure. We will explain how this structure provided an intuitive guideline for the study of a family of reaction-diffusion and kinetic equations. The solutions to the evolution equations will correspond to the systems of rays that are calibrated with respect to a functional in the space of positive measures that will act as a KAM function.

1.1.5 Minimizing movement scheme for calibrated curves

In this section we show how one can use an analogue of the method of Section 1.1 to build solutions to

$$\begin{cases} \dot{x}_t = \mathcal{L}_{x_t}(-d_{x_t}E) \quad \forall t \in [0,\infty), \\ x(0) = x_0. \end{cases}$$

Here, E and L are given and satisfy the assumptions of Section 1.1 and 1.4. Also, the Legendre transform $\mathcal{L}_{x_t}(-d_{x_t}E)$ is defined as in (1.1.16).

We proceed as in Section 1.1. Given a time step $\tau > 0$, we produce a sequence of points $\{x_{k\tau}\}_{k=0}^{\infty}$ where $x_{k\tau}$ is determined from $x_{(k-1)\tau}$ by setting

$$x_{k\tau} = \mathrm{arming}_x h_\tau(x_{(k-1)\tau}, x) + E(x),$$

where

$$h_t(x,y) = \inf_{\alpha} \int_0^t L(\alpha_s, \dot{\alpha}_s)\, ds.$$

Here, α ranges along all curves satisfying $\alpha(0) = x$ and $\alpha(t) = y$.

Let us suppose $\alpha : [0,\tau) \to \mathbb{R}^d$ is a curve satisfying $\alpha(0) = x_{(k-1)\tau}$, $\alpha(\tau) = x_{k\tau}$, and

$$\int_0^\tau L(\alpha, \dot{\alpha})dt + E(\alpha(\tau) = \inf_x h_\tau(x_{(k-1)\tau}, x) + E(x). \tag{1.1.17}$$

Then, we can perturb α by deforming it in a way that fixes the end points. That is is, we consider a map $\tilde{\alpha} : [0,\tau) \times (-\varepsilon,\varepsilon) \to \mathbb{R}^d$ such that

$$\begin{cases} \tilde{\alpha}(t,0) = \alpha(t) \quad \forall t \in [0,\tau), \\ \tilde{\alpha}(0,s) = \alpha(0) \quad \forall s \in (-\varepsilon,\varepsilon), \\ \tilde{\alpha}(\tau,s) = \alpha(\tau) \quad \forall s \in (-\varepsilon,\varepsilon). \end{cases} \tag{1.1.18}$$

Then, by (1.1.17), we must have

$$
\begin{aligned}
0 &= \frac{d}{ds}\bigg|_{s=0} \int_0^\tau L\big(\tilde{\alpha}(t,s), \partial_t \tilde{\alpha}(t,s)\big)\, dt \\
&= \int_0^\tau \nabla_x L(\tilde{\alpha}(t,0), \partial_t \tilde{\alpha}(t,0)\big)\partial_s \tilde{\alpha}(t,0) \\
&+ \nabla_v L\big(\tilde{\alpha}(t,0), \partial_t \tilde{\alpha}(t,0)\big)\partial_s \partial_t \tilde{\alpha}(t,0)\, dt \\
&= \int_0^\tau \Big[\nabla_x L(\alpha, \dot{\alpha}) - \frac{d}{dt}\nabla_v L(\alpha, \dot{\alpha})\Big]\partial_s \tilde{\alpha}(t,0)\, dt,
\end{aligned}
$$

where in the last equality we have used integration by parts and conditions (1.1.18). Since, $\partial_s \tilde{\alpha}(t,0)$ can be varied arbitrarily for any t in $(0,\tau)$, we conclude that

$$
\frac{d}{dt}\nabla_v L(\alpha, \dot{\alpha}) = \nabla_x L(\alpha, \dot{\alpha}). \tag{1.1.19}
$$

This condition is called the Euler Lagrange equation and is satisfied by extremals of the action.

Now, we can vary α in such a way that $\alpha(\tau)$ is free. That is. we consider a map $\tilde{\alpha} \,:\, [0,\tau) \times (-\varepsilon, \varepsilon) \to \mathbb{R}^d$ satisfying

$$
\begin{cases}
\tilde{\alpha}(t,0) = \alpha(t) & \forall t \in [0,\tau), \\
\tilde{\alpha}(0,s) = \alpha(0) & \forall s \in (-\varepsilon, \varepsilon).
\end{cases} \tag{1.1.20}
$$

Then, by (1.1.17), we must have

$$
\begin{aligned}
0 &= \frac{d}{ds}\bigg|_{s=0} \int_0^\tau L\big(\tilde{\alpha}(t,s), \partial_t \tilde{\alpha}(t,s)\big)\, dt + E(\tilde{\alpha}(\tau, s)) \\
&= \int_0^\tau \nabla_x L(\tilde{\alpha}(t,0), \partial_t \tilde{\alpha}(t,0)\big)\partial_s \tilde{\alpha}(t,0) \\
&+ \nabla_v L\big(\tilde{\alpha}(t,0), \partial_t \tilde{\alpha}(t,0)\big)\partial_s \partial_t \tilde{\alpha}(t,0)\, dt \\
&= \int_0^\tau \big[\nabla_x L(\alpha, \dot{\alpha}) - \frac{d}{dt}\nabla_v L(\alpha, \dot{\alpha})\Big]\partial_s \tilde{\alpha}(t,0)\, dt \\
&+ [\nabla_v L(\alpha(\tau), \dot{\alpha}(\tau)) + \nabla_{\alpha(\tau)} E]\, \partial_s \tilde{\alpha}(\tau, 0) \\
&= [\nabla_v L(\alpha(\tau), \dot{\alpha}(\tau)) + \nabla_{\alpha(\tau)} E]\, \partial_s \tilde{\alpha}(\tau, 0).
\end{aligned}
$$

Here, in the last two equalities we have used integration by parts and conditions (1.1.19) and (1.1.20) . Hence, since $\partial_s \tilde{\alpha}(\tau, 0)$ can be varied arbitrarily, we must have

$$
0 = \nabla_v L(\alpha(\tau), \dot{\alpha}(\tau)) + \nabla_{\alpha(\tau)} E.
$$

By (1.1.16), this is equivalent to the fact that

$$
\dot{\alpha}(\tau) = \mathcal{L}_{x_{k\tau}}(-d_{x_{k\tau}} E),
$$

and is the analogue of condition (1.1.3).

In [42] Figalli, Gangbo, and Yolcu follow the analogue of these ideas in $(\mathbb{P}_2(\mathbb{R}), W_2)$ to build solution to a family of non-linear parabolic equations. Moreover, in Sections 5 and 6 and in Chapter 2, we use these ideas at the level of action functionals in the space of positive path of measures to build solutions to a family of reaction-diffusion equations with boundary conditions and kinetic equations.

1.1.6 Dynamics for systems of rays

In our previous discussion, we saw that gradient flows are a particular case of weak KAM structures where at each point we have a calibrated curve related to the KAM function. Under compactness assumptions, as in the case of the gradient flows, when we let $t \to -\infty$ the calibrated curves converge to the so-called Aubry set, which plays crucial role in the long term dynamics (see [88, Proposition 4.1]). This set is defined by

$$\mathcal{A}(H) := \big\{x \in M \ : \ h(x,x) = 0\big\},$$

where,

$$h(x,y) := \liminf_{t\to\infty} \big\{h_t(x,y) + c[H]t\big\},$$

and

$$h_t(x,y) = \inf_{\alpha(0)=1,\alpha(t)=y} \int_0^t L(\alpha(s), \dot{\alpha}(s)ds.$$

Here, H is the Hamiltonian induced by L via the Legendre transform, and $c[H]$ is the critical value introduced in section 1.4.

Now, we state a celebrated conjecture of Mañe that describes the structure of this set in a compact Riemannian manifold (M, g) for generic strictly convex and super-linear Hamiltonians in $C^k(M)$.

Conjecture (Mañe conjecture). *Let $H : T^*M \to \mathbb{R}$ be a Hamiltonian satisfying:*

(i) H is in $C^k(M)$.

(ii) For every $K \geq 0$, there is a finite constant $C^*(K)$ such that,

$$H(x,p) \geq K||p||_x + C^*(K).$$

(iii) For every $(x,p) \in T^*M$, the second derivative along the fibers $\frac{\partial^2 H}{\partial p^2}(x,p)$ is positive definite.

Then, exists a dense set D in $C^k(M)$ such that, for every $V \in D$, the Aubry set of the Hamiltonian, $H+V$ is either an equilibrium point or a periodic orbit.

In Section 5 and in Chapter 5 we use the relationship between calibrated curves and the Aubry set as an intuitive guideline to obtain our emergence of phase concentration result for the Kuramoto Sakaguchi equation.

1.2 Otto calculus: The formal Riemannian structure

In this section, we describe the formal Riemannian structure of $(\mathbb{P}_2(M), W_2)$, introduced by Felix Otto. We begin by recalling the existence of metric derivatives for absolutely continuous curves in metric spaces.

1.2.1 Absolutely continuous curves

Let (Y,d) be a metric space and let $y : [0,1] \to Y$ be a curve. Then we say that y is absolutely continuous if there exists a function f in $L^1(0,1)$

such that

$$d(y(t), y(s)) \leq \int_t^s f(r)\, dr, \quad \forall t < s \in [0,1].$$

Whenever the above condition is satisfied, we have that for almost every t the metric derivative $|\dot{y}|$ exists and is given by

$$|\dot{y}(t)| := \lim_{h \to 0} \frac{d(y_{t+h}, y_t)}{|h|}.$$

The first step towards the formal Riemannian structure of $(\mathbb{P}_2(M), W_2)$ is understanding the relationship between vector fields and absolute continuous curves in such a metric space.

Theorem 1.2.1. *Let M be a smooth complete Riemannian manifold without boundary. Then the following holds:*

(i) For every absolutely continuous curve $t \to \mu_t \subset \mathbb{P}_2(M)$ there exists a Borel family of vector fields v_t on M such that $||v_t||_{L^2(\mu_t)} \leq |\dot{\mu}_t|$ for a.e t and

$$\frac{d}{dt}\mu_t + \nabla \cdot (v_t \mu_t) = 0, \tag{1.2.21}$$

in the sense of distributions.

(ii) If (μ_t, v_t) satisfies the above equation in the sense of distributions and $\int_0^1 \int |v_t|^2\, d\mu_t dt < \infty$, then up to redefining $t \to \mu_t$ on a negligible set of times, (μ_t) is an absolutely continuous curve on $\mathbb{P}_2(M)$ and $|\dot{\mu}_t| \leq ||v_t||_{L^2(\mu_t)}$, for a.e $t \in [0,1]$.

Here, we are using the definition

$$||W||_{L^2(\mu)} := \left(\int |W|^2 \, d\mu \right)^{\frac{1}{2}}, \tag{1.2.22}$$

for any vector field W and any probabilty measure μ. We refer the reader to [5, Theorem 2.29] for the proof of this theorem.

1.2.2 The tangent plane

Given an absolutely continuous curve $t \to \mu_t$, the choice of the vector field v_t is not unique.

In a Riemannian manifold (M, g) one can define the distance between two points as the square root of the minimum among of kinetic energy needed by a curve that joins them. Hence, given x and y in M we have

$$d(x, y) = \sqrt{\inf_{\gamma} \int_0^1 g(\dot{\gamma}, \dot{\gamma}) \, ds}, \tag{1.2.23}$$

where γ ranges among all curves such that $\gamma(0) = x$ and $\gamma(1) = y$. Additionally, g is the metric tensor in M. For notational simplicity, sometimes we will omit the metric tensor so that we will denote $g(W, W)$ by $|W|^2$ or $g(A, W)$ by $\langle A, W \rangle$ or just AW.

The analogue of identity (1.2.23) in $(\mathbb{P}_2(M), W_2)$ is given by the Benamou-Brenier formula

$$W_2(\mu, \nu) := \sqrt{\min_{v_t, \mu_t} \int_0^1 \int |v_t|^2 \, d\mu_t dt},$$

where the minimum is taken over all absolutely continuous curves $t \to \mu_t$ and time-dependent vector fields $t \to v_t$ satisfying (1.2.21) and such that $\mu_0 = \mu$ and $\mu_1 = \nu$.

This strongly suggests that at each μ the metric tensor in $(\mathbb{P}_2(M), W_2)$ should be given by:

$$\langle A, W \rangle_\mu = \int g(A, W)\, d\mu.$$

Such a metric tensor provides a way of canonically selecting the family of vector fields associated with a given absolutely continuous curve in $(\mathbb{P}_2(M), W_2)$. Indeed, since the constraint

$$\frac{d}{dt} \int \zeta \, d\mu_t = \int \nabla \zeta v_t \, d\mu_t \quad \forall \zeta \in C^\infty(M), \tag{1.2.24}$$

is linear and the map

$$v_t \to \int |v_t|^2 d\mu_t,$$

is strictly convex, for any t there exists a unique minimizer v_t satisfying (1.2.24).

By perturbing such a functional among vector fields satisfying (1.2.24), one finds that the minimizer satisfies

$$\int |v_t|^2 \, d\mu_t < \infty,$$

and

$$\left\{ \int v_t W \, d\mu_t = 0, \quad \forall W \quad \text{s.t} \quad \int \nabla \zeta W \, d\mu_t = 0 \quad \forall \zeta \in C^\infty(M) \right\}.$$

This set actually agrees with the $L^2(\mu_t)$ closure of

$$\{\nabla\varphi \,:\, \varphi \in C_c^\infty(M)\}$$

(see (1.2.22)).

This motivates the following definition:

Definition: Let μ in $\mathbb{P}_2(M)$. Then, the tangent space $T_\mu\mathbb{P}_2(M)$ is defined as the $L^2(\mu)$ closure of the set

$$\{\nabla\varphi \,:\, \varphi \in C_c^\infty(M)\}.$$

The inner product

$$\langle \nabla\varphi, \nabla\varphi' \rangle_\mu = \int \langle \nabla\varphi, \nabla\varphi \rangle \, d\mu,$$

turns this space into a Hilbert space.

1.2.3 The velocity field along a curve and the geodesics equation

Henceforth, we will say that $t \to \mu_t$ is a curve with velocity field $t \to \nabla\varphi_t$, whenever

$$\frac{d}{dt}\int \zeta \, d\mu_t = \int \langle \nabla\zeta, \nabla\varphi_t \rangle \, d\mu_t \quad \forall \zeta \in C^\infty(M) \quad \text{and} \quad \forall t \in [0, \infty).$$

In a Riemannian Manifold, the geodesic equation is given by:

$$D_{\dot\gamma}\dot\gamma = 0,$$

where $D_X Y$ is the Levi-Civita connection in the direction of the vector field X applied to the vector field Y.

Moreover, for any curve γ, we have that

$$D_{\dot\gamma}\langle X, Y\rangle = \langle D_{\dot\gamma}X, Y\rangle + \langle X, D_{\dot\gamma}Y\rangle.$$

In particular,

$$D_{\dot\gamma}\frac{1}{2}\langle\dot\gamma,\dot\gamma\rangle = \langle D_{\dot\gamma}\dot\gamma,\dot\gamma\rangle.$$

On the other hand, if $t \to \mu_t$ is a curve with velocity field $t \to \nabla\varphi_t$, then we have

$$\frac{d}{dt}\int \frac{1}{2}|\nabla\varphi_t|^2\, d\mu_t = \int \nabla\Big(\partial_t\varphi + \frac{1}{2}|\nabla\varphi_t|^2\Big)\nabla\varphi_t\, d\mu_t.$$

which suggests that the geodesic equation in $(\mathbb{P}_2(M), W_2)$ should be given by

$$\partial_t\varphi + \frac{1}{2}|\nabla\varphi_t|^2 = 0.$$

1.2.4 The gradient of the relative entropy

Now, we study the gradient flow of the relative entropy in the formal Riemannian structure of $(\mathbb{P}_2(M), W_2)$. For this purpose we consider the functional

$$H(\mu|e^{-V}) = \begin{cases} \int_M \rho\log\rho + V\rho\, dx, & \text{if} \quad \mu{=}\rho\ \ dx,\\ \infty & \text{otherwise.}\end{cases}$$

Here, e^{-V} is a probability density in $\mathbb{P}_2(M)$ and V is in $C^\infty(M)$.

Recall that if $\gamma : (-\varepsilon, \varepsilon) \to \mathbb{R}$ and $E : \mathbb{R} \to \mathbb{R}$ are differentiable, then

$$\frac{d}{dt}\bigg|_{t=0} E(\gamma) = \langle \dot{\gamma}(0), \nabla_{\gamma(0)} E \rangle.$$

Let $t \to \mu_t$ be a curve with velocity field $t \to \nabla\varphi_t$ satisfying

$$\mu_0 = \rho,$$

and

$$\varphi_0 = \varphi.$$

Then,

$$\frac{d}{dt}\bigg|_{t=0} H(\mu_t | e^{-V}) = \int (\log \rho + V) \partial_t \rho \, dx = \int \nabla\varphi \nabla(\log \rho + V) \rho \, dx.$$

Since

$$\nabla(\log \rho + V) \in T_\rho \mathbb{P}_2(M),$$

this computation strongly suggests that $\nabla(\log \rho + V)$ should be the gradient of the functional $\rho \to H(\rho | e^{-V})$ at ρ.

Moreover, since

$$\langle \nabla(\log \rho + V), \nabla(\log \rho + V) \rangle_\rho = \int |\nabla \log \frac{\rho}{e^{-V}}|^2 \rho \, dx,$$

it will be convinient to define

$$I(\mu | e^{-V}) = \begin{cases} \int_M |\nabla \log \frac{\rho}{e^{-V}}|^2 \rho \, dx, & \text{if} \quad \mu = \rho \ \ dx, \\ \infty & \text{otherwise.} \end{cases}$$

1.2.5 The Hessian of the relative entropy

If γ satisfies $\ddot{\gamma} = 0$ (the geodesic equation in $\mathbb{R}^d$) and $D^2 E \geq \lambda\,\mathrm{Id}$, then we have

$$\begin{aligned}\frac{d^2}{dt^2}E(\gamma) &= \frac{d}{dt}\langle \nabla_\gamma E, \dot{\gamma}\rangle \\ &= \langle \ddot{\gamma}, \nabla_\gamma E\rangle + \langle \dot{\gamma}, D^2_\gamma E \dot{\gamma}\rangle \\ &= \langle D^2_\gamma E \dot{\gamma}, \dot{\gamma}\rangle \\ &\geq \lambda |\dot{\gamma}|^2.\end{aligned}$$

On the other hand, if

$$\mathrm{Ric}_M + D^2 V \geq \lambda\,\mathrm{Id}, \tag{1.2.25}$$

and $t \to \rho_t$ is a curve with velocity field $t \to \nabla\varphi_t$ satisfying the geodesic equation in $(\mathbb{P}_2(M), W_2)$

$$\partial_t \varphi_t + \frac{1}{2}|\nabla\varphi_t|^2 = 0,$$

then we have

$$\begin{aligned}\frac{d^2}{dt^2}H(\mu_t|e^{-V}) &= \frac{d}{dt}\int \nabla\varphi_t \nabla(\log\rho - V)\rho_t\,dx \\ &= \int \mathrm{tr}([D^2\varphi_t]^T D^2\varphi_t)\rho_t\,dx + \int \langle(\mathrm{Ric}_M + D^2 V)\nabla\varphi_t, \nabla\varphi_t\rangle \rho_t\,dx \\ &\geq \lambda \int \langle \nabla\varphi_t, \nabla\varphi_t\rangle \rho_t\,dx,\end{aligned}$$

(see [84, Section 3]).

Thus, it is reasonable to expect some form of λ-convexity for the functional $\mu \to H(\mu|e^{-V})$, whenever $\mathrm{Ric}_M + D^2 V \geq \lambda\,\mathrm{Id}$. Indeed, this notion is called displacement convexity; we refer the reader to [5] for an introduction to this topic.

1.2.6 The gradient flow of the relative entropy

If we compare the definition of a gradient flow in $\mathbb{R}^d$ with the relationship between absolutely continuous curves in $(\mathbb{P}_2(M), W_2)$ described in Section 2.1, then it is natural to define the gradient flow of $\mathcal{E}$ starting at ρ as a curve $t \to \rho_t$ with velocity field $t \to v_t$ such that

$$v_t = -\nabla(\log \rho_t + V) \quad \forall t \in [0, \infty),$$

(see section 2.4).

This is a equivalent to the fact that ρ_t is a weak solution of the Fokker-Planck equation

$$\partial_t \rho_t = \mathrm{div}(\nabla \rho_t - \rho_t \nabla V).$$

In [84], for M compact, Otto and Villani use (1.2.25) to prove the inequalities

$$H(\mu|e^{-V}) \leq \frac{1}{2\lambda} I(\mu|e^{-V}), \tag{1.2.26}$$

and

$$W_2(\mu, e^{-V}) \leq \sqrt{\frac{2}{\lambda} H(\mu|e^{-V})}, \tag{1.2.27}$$

which are the analogue of inequalities (1.1.8) and (1.1.12) for the gradient flow of the relative entropy in the formal Riemannian structure of $(\mathbb{P}_2(M), W_2)$.

Additionally, by using the work of Holley and Stroock [63], Otto and Villani show that if $\tilde{V} = V + \psi$ with ψ in $L^\infty(M)$, and $e^{-\tilde{V}}$ is a probability density, then

$$H(\mu|e^{-V}) \leq \frac{1}{2\tilde{\lambda}} I(\mu|e^{-V}), \tag{1.2.28}$$

where $\tilde{\lambda} := \lambda e^{-\operatorname{osc}(\psi)}$ and $\operatorname{osc}(\psi) = \sup \psi - \inf \psi$.

1.3 The path space of a smooth Manifold.

In this section, we review the formal differential structure of the space of paths introduce by Milnor in [78, Part III]. This structure is the essential tool that allow us to carry formal computations in the weak KAM structures developed to study kinetic and reaction-diffusion equations with boundary conditions in Sections 5 and 6.

Given a smooth Manifold M, we denote by $\Omega(M)$ the space of all smooth maps from $[0,1]$ to M. Here, we give $\Omega(M)$ the formal structure of a infinite dimensional manifold.

We define the tangent space of Ω at a path ω to be the vector space consisting of all smooth vector fields W along ω such that $W(0) = 0$ and $W(1) = 0$.

Such a space will be denoted by $T_\omega\Omega$. Given a real valued function F on Ω it is natural to ask what the differential of F at a given path is, which we shall denote by $d_\omega F$. Such a differential is linear map from $T_\omega\Omega$ to $\mathbb{R}$. When F is a differentiable function on a smooth manifold M and X_m is a vector in T_mM for some m in M, then

$$d_mF(X_m) = \frac{d}{dt}\bigg|_{t=0} F(\alpha),$$

where $\alpha : (-\varepsilon, \varepsilon) \to M$ satisfies

$$\alpha(0) = m,$$

and

$$\frac{d\alpha}{dt}\bigg|_{t=0} = X_m.$$

To carry out the analogous construction for $F : \Omega \to \mathbb{R}$, given a path ω we define a variation as a map $\alpha : [0,1] \times (-\varepsilon, \varepsilon) \to M$, for some $\varepsilon > 0$, such that

$$\begin{cases} \alpha(t,0) = \omega(t) & \forall t \in [0,\tau), \\ \alpha(0,s) = \omega(0) & \forall s \in (-\varepsilon,\varepsilon), \\ \alpha(1,s) = \omega(1) & \forall s \in (-\varepsilon,\varepsilon). \end{cases}$$

The map $s \to \alpha(\cdot, s)$ can be interpreted as a smooth path in Ω. We define its velocity vector as the element in $T_\omega\Omega$ given by

$$W_t = \partial_s \alpha(t, 0).$$

Given W in $T_\omega\Omega$, it is always possible to find a variation that has such a velocity vector (see [78, Part III]). Hence, one can formally define

$$d_\omega F(X_\omega) = \frac{d}{dt}\bigg|_{t=0} F(\alpha),$$

where α is a variation of ω corresponding to X_ω in $T_\omega\Omega$.

Let us consider the following example. We define the energy functional in the space of paths between x and y as

$$E(\gamma) := \frac{1}{2}\int_0^1 \langle \dot{\gamma}, \dot{\gamma}\rangle \, dt,$$

for any map $\gamma \,:\, [0,1] \to M$ such that $\gamma(0) = x$ and $\gamma(1) = y$.

If α is a variation of γ with velocity vector X in $T_\gamma\Omega$, then we have

$$\frac{d}{ds}\bigg|_{s=0} E(\alpha) = -\int_0^1 \langle X, D_{\dot{\gamma}}\dot{\gamma}\rangle \, dt,$$

(see [78, Theorem 12.2]).

Hence, one sees that geodesics are critical points of E.

1.4 The homogeneous Vicsek model as a gradient flow

In this section, we discuss techniques that can be applied to gradient flows that do not possess uniform convexity lower bounds. We use these methods to obtain convergence rates and stability estimates for the homogeneous Vicsek model:

$$\begin{cases} \partial_t \rho = \Delta_\omega \rho - \nabla_\omega \cdot \Big(\rho \mathbb{P}_{\omega^\perp}\Omega_\rho\Big) \quad \text{in} \quad \mathbb{S}^{d-1}, \\ \Omega_\rho = \frac{J_\rho}{|J_\rho|}, \quad J_\rho = \int_{\mathbb{S}^{d-1}} \omega\rho \; d\omega, \\ \rho(0) = \rho_0 \quad \text{in} \quad \mathbb{S}^{d-1}. \end{cases} \tag{1.4.29}$$

Here, the operators ∇_ω and Δ_ω denote the gradient and the Laplace-Beltrami operator on the sphere $\mathbb{S}^{d-1}$. The density function $\rho(t,\omega)$ is a one-particle distribution at time t with direction $\omega \in \mathbb{S}^{d-1}$. The term $\mathbb{P}_{\omega^\perp}\Omega$ denotes the projection of the vector Ω onto the normal plane to ω, describing the mean-field force that governs the orientational interaction of self-driven particles by aligning them with the direction Ω determined by the flux J.

The equation (1.4.29) was formally derived by Degond and Motsch [31] as a mean-field limit of the discrete Vicsek model [3, 15, 26, 50] with stochastic dynamics. Recently, the stochastic Vicsek model has received extensive attention in the mathematical topics such as the mean-field limit, hydrodynamic limit, and phase transition. Bolley, Canizo, and Carrillo [16] have rigorously justified the mean-field limit when the unit vector Ω in the force term of [31] is replaced by a more regular-vector. Additionally, Degond, Frouvelle, and Liu [28] provided a complete and rigorous description of phase transitions. However, in their work, in order to overcome the difficulty that (1.4.29) is not defined when $J = 0$, Ω is replaced by $\nu(|J|)\Omega$ and there is a noise intensity $\tau(|J|)$ in front of Δ_ω, where the functions ν and τ satisfy

$$|J| \to \frac{\nu(|J|)}{|J|}, \quad |J| \to \tau(|J|), \quad \text{and} \quad |J| \to \tau(|J|) \quad \text{are Lipschitz and bounded.}$$

More recently, Gamba and Kang [49] proved the existence and uniqueness of weak solutions to the kinetic Kolmogorov-Vicsek model with the singular force field $\mathbb{P}_{\omega^\perp}\Omega$, under the a priori assumption that $|J| > 0$. The solution constructed in [49] are in L^∞ in time and $L^p(D)$, where D is both in x-space and v-space. (and so they are weak solutions in the classical sense) and have stability with respect to the initial data, under the a priori assumption of $|J| > c$ uniformly in time and space. See statement of Theorem 2.1 in [49] equation (2.3) to (2.7). This statement also holds for the space-homogeneous setting under the a priori assumption of $|J| > 0$.

In Chapter 3, in collaboration with Alessio Figalli and Moon-Jin Kang, we obtain well-posedness, stability, and convergence to equilibrium for (1.4.29), when the initial data satisfies that $|J_{\rho_0}| > 0$ and $\mathcal{E}(\rho_0) < \infty$. Here, the functional $\mathcal{E}$ is defined by

$$\mathcal{E}(\rho) := \int \rho \log \rho - |J_\rho|.$$

We show that (1.4.29) can be seen the gradient flow of $\mathcal{E}$ in $(\mathbb{P}_2(\mathbb{S}^{d-1}), W_2)$. Indeed, for any curve $s \to \rho_s$ with velocity field $t \to \nabla\varphi_s$ we have,

$$\begin{aligned}\frac{d}{ds}\mathcal{E}(\rho_s) &= \int_{S^{d-1}} \nabla_\omega\varphi[\nabla_\omega \log \rho_s - \nabla_\omega(\omega \cdot \Omega_{\rho_s})]\,\rho_s\, d\omega \\ &= \langle \nabla\varphi, \nabla_\omega \log \rho_s - \nabla_\omega(\omega\cdot\Omega_{\rho_s})\rangle_{\rho_s}.\end{aligned}$$

Consequently, by the discussion in Section 2 we see that $\nabla_\omega \log \rho_s - \nabla_\omega(\omega \cdot \Omega_{\rho_s}) \in T_\rho\mathbb{P}_2(\mathbb{S}^{d-1})$, can be regarded as the gradient of $\mathcal{E}$ at ρ.

By definition, a curve $t \to \rho_t$ with velocity field $-\big[\nabla_\omega \log \rho_t - \nabla_\omega(\omega \cdot \Omega_{\rho_t})]$ satisfies

$$\frac{d}{dt}\int \zeta\rho_t\, dx = -\int_{\mathbb{S}^{d-1}} \langle \nabla\zeta, \nabla_\omega \log\rho_t - \nabla_\omega(\omega\cdot\Omega_{\rho_t})\rangle \rho_t\, dx \quad \forall \zeta \in C^\infty(\mathbb{S}^{d-1}) \quad \text{and} \quad \forall t \in [0,\infty),$$

which formally corresponds to a solution of (1.4.29) .

To overcome the difficulty that the equation is not defined when $|J_\rho| = 0$, we regularize the equation. We do this by considering the gradient flow of

$$\int \rho \log \rho - \sqrt{|J_\rho|^2 + \varepsilon}.$$

Such a gradient flow solves (1.4.29) where Ω_ρ is replaced by

$$\Omega_\rho^\varepsilon := \frac{J_\rho}{\sqrt{|J_\rho|^2 + \varepsilon}}.$$

We show the existence of the regularized equation by using the Jordan-Kinderlehrer-Otto scheme [66]. Then, we get the estimate

$$|J_{\rho^\varepsilon(t)}|^2 \geq |J_{\rho_0}|^2 e^{-2(d-1)t}, \tag{1.4.30}$$

which uniform in ε (see Lemma 3.3 from Chapter 3). Here, ρ^ε is a solution to the regularized equation.

Finally, we use (1.4.30) together with gradient structure to show equicontinuity of the regularizations on the Wasserstein metric which allows us to pass to the limit as $\varepsilon \to 0$ and get the existence of solutions to (1.4.29).

(It is worth noticing that the existence of solutions to the regularized equation could also be proved by using the techniques in [49]. However, we used the Jordan-Kinderlehrer-Otto scheme because it gave us useful estimates for the limiting system.)

Moreover, if the velocity field $s \to \nabla\varphi_s$ satisfies the geodesic equation in $\mathbb{P}_2(\mathbb{S}^{d-1})$

$$\partial_s\varphi_s + \frac{1}{2}|\nabla\varphi_s|^2 = 0,$$

then

$$\begin{aligned}\frac{d^2}{dt^2}\bigg|_{t=0}\mathcal{E}(\rho_s) &= \int_{\mathbb{S}^{d-1}} \mathrm{tr}([D^2\varphi_s]^T D^2\varphi_s)\,\rho_s\,d\omega + (d-2)\int_{\mathbb{S}^{d-1}} |\nabla\varphi_s|^2\rho_s\,d\omega \\ &+ \int_{\mathbb{S}^{d-1}} \nabla\varphi_s\,D^2(\Omega_\rho\cdot\omega)\,\nabla\varphi_s\,\rho_s\,d\omega \\ &- \frac{1}{\sqrt{|J_{\rho_s}|^2}}\left(\Big|\int_{\mathbb{S}^{d-1}} \nabla\varphi_s\,\rho_s\,d\omega\Big|^2 - \Big(\int_{\mathbb{S}^{d-1}} \Omega_{\rho_s}\cdot\nabla\varphi_s\,\rho_s\,d\omega\Big)^2\right) \\ &\geq -\frac{1}{|J_{\rho_s}|}\int_{\mathbb{S}^{d-1}} |\nabla\varphi_s|^2\rho_s\,d\omega,\end{aligned}$$

(see Lemma 3.1 from Chapter 3).

As in Section 2.5, this condition suggests a pointwise lower bound on the Wasserstein Hessian of $\mathcal{E}$ at ρ given by $-|J_\rho|^{-1}$. For gradient flows in $\mathbb{R}^d$ with uniform Hessian lower bound, this condition is enough to give stability of solutions (see Section 1.3).

Indeed, to obtain the short time stability estimate, we show a lower bound on $|J_\rho|$ along solutions and we prove that the map $\rho \to |J_\rho|$ is continuous in the Wasserestein metric. Then, we perform the analogue of the estimate in Section 1.3, where the Wasserstein geodesic between two initial conditions ρ_0 and $\tilde{\rho}_0$

plays the role of the line $s \to x + s(\tilde{x} - x)$. By propagating the lower bound on $|J_\rho|$ along such a Wassrestein geodesic we obtain a local version of the displacement convexity condition and we achive stability for solutions to (1.4.29).

Additionally, we show that for $d \geq 3$ any solution of (1.4.29) with initial finite logarithmic entropy and nonvanishing momentum converges exponentially (with explicit constant) to a Fisher-Von Mises distribution: for any given $\Omega \in \mathbb{S}^{d-1}$, these are given by

$$M_\Omega(\omega) = \frac{e^{\omega\cdot\Omega}}{\int_{\mathbb{S}^{d-1}} e^{\omega\cdot\Omega}\, d\omega}.$$

We prove convergence to equilibrium, for $d \geq 3$, by using the fact that (1.4.29) is a gradient flow of the functional:

$$\mathcal{E}(\rho) = \int_{\mathbb{S}^{d-1}} \log\left(\frac{\rho}{M_{\Omega_\rho}}\right)\rho\, d\omega = \inf_{\Omega\in\mathbb{S}^{d-1}} \int_{\mathbb{S}^{d-1}} \log\left(\frac{\rho}{M_\Omega}\right)\rho\, d\omega. \tag{1.4.31}$$

As suggested by the above expression, at each ρ, $\mathcal{E}(\rho)$ concides at first order with the relative entropy with respect to M_{Ω_ρ}. Indeed,

$$\frac{d}{dt}\mathcal{E}(\rho_t) = \frac{d}{dt}H(\rho_t|M_{\Omega_{\rho_t}}) = I(\rho_t|M_{\Omega_{\rho_t}}). \tag{1.4.32}$$

For $d \geq 3$ this relative entropy is a bounded perturbation (in the Holley and Stroock sense [63]) of a strictly convex functional on $(\mathbb{P}_2(\mathbb{S}^{d-1}), W_2)$. Hence, the relative entropy satisfies the logarithmic Sobolev inequality (see Section 2.4). By using this, we show that $\mathcal{E}$ decays exponentially.

The method we used in [44] does not work to show convergence to equilibrium of (1.4.29) in the case $d = 2$. The reason is that the celebrated Bakry-Emery condition for the logarithmic Sobolev inequality is not available for that dimension.

In Chapter 4, in collaboration with Moon-Jin Kang, we obtain exponential convergence to equilibrium (with explicit rate) for (1.4.29) when $d = 2$. In this case, $\mathcal{E}$ does not agree at first order with a bounded perturbation of a displacement convex function and hence quantitative convergence rates cannot be obtained by using the existing theory [6]. To overcome this difficulty we developed a general method that relies on two central estimates.

The first one is a localized version of the logarithmic Sobolev inequality, which we prove by linearization. Such an inequality is given by

$$\int \log\left(\frac{\rho}{e^{-\Psi}}\right)\rho \, d\omega \leq \frac{2\pi^2 e^{2||\Psi||_\infty}(1+\frac{1}{15}\varepsilon_*)}{1-\frac{7}{6}\varepsilon_*}\int |\nabla \log \frac{\rho}{e^{-\Psi}}|^2 \rho \, d\omega.$$

The above estimate holds for any ρ and $e^{-\Psi}$ in $\mathbb{P}_2(\mathbb{S}^1)$ such that Ψ and ρ belong to $C^1(\mathbb{S}^1)$ and $||\rho - e^{-\Psi}||_{L^\infty(\mathbb{S}^1)} \leq \varepsilon_* \leq 1/10$.

The second one is a global argument that exploits the gradient flow structure. There, inspired by the expression for the Wasserstein Hessian we computed in [44], we were able to obtain a differential inequality on the Wasserstein slope.

From (1.4.32), we see that $I(\rho|M_{\Omega_\rho})$ plays the role of $|\nabla_x E|^2$ for the gradient flow of $\mathcal{E}$ in the formalism introduced in Section 2. In Section 1.2, under the assumption that $D^2 E \geq \lambda\,\mathrm{Id}$, we proved the inequality

$$\frac{d}{dt}|\nabla_x E|^2 \leq -\lambda |\nabla_x E|^2.$$

Inspired by this, by direct computation in Lemma 3.1 we show that

$$\frac{d}{dt} I(\rho_t|M_{\Omega_{\rho_t}}) \leq 2\left(1 + \frac{1}{|J_{\rho_t}|}\right) I(\rho|M_{\Omega_\rho}). \tag{1.4.33}$$

Consequently, for every $T \geq 0$, we have

$$I(\rho_t|M_{\Omega_{\rho_t}}) \leq e^{C\frac{2}{\min_{t\in[T,t]}|J_{\rho_t}|}(t-T)} I(\rho_T|M_{\Omega_{\rho_T}}) \quad t \geq T,$$

for some $C > 0$.

Using the previous estimates and (1.4.30), we can find a time T_0 depending explicitly on $\mathcal{E}(\rho_0)$ and $|J_{\rho_0}|$ such that

$$\mathcal{E}(\rho_t) \leq C_* I(\rho_t|M_{\Omega_{\rho_t}}) \leq C_* \varepsilon_* \quad \text{in} \quad [T_0, \infty),$$

(see Section 3.1 of Chapter 4).

Hence, from (1.4.32) and (1.4.33), we get

$$\begin{aligned}
\mathcal{E}(\rho_t) &= \mathcal{E}(\rho_t) - \mathcal{E}(\rho_{T_0}) + \mathcal{E}(\rho_{T_0}) \\
&= \int_t^{T_0} I(\rho_s|M_{\Omega_{\rho_s}})ds + C_* I(\rho_{T_0}|M_{\Omega_{\rho_{T_0}}}) \\
&\leq I(\rho_t|M_{\Omega_{\rho_t}})\Bigg[e^{(t-T_0)\min\frac{2}{\min_{t\in[T,t]}|J_{\rho_t}|}(t-T)} \\
&\qquad + \int_t^{T_0} e^{(s-T_0)\min\frac{2}{\min_{t\in[T,t]}|J_{\rho_t}|}(t-T)} ds\Bigg].
\end{aligned}$$

When we combine this with the estimate on the momentum we get an inequality of the form

$$\mathcal{E}(\rho_t) \leq B_{\rho_0} I(\rho_t|M_{\Omega_{\rho_t}}),$$

(where the dependence of B_{ρ_0} is explicitly quantified in Chapter 4). By using this inequality in the same way inequality (1.1.8) was used in Section 1.2, we get the quantitative convergence rate.

1.5 Formal weak KAM structure for kinetic equations and the Kuramoto Sakaguchi model

In this section,, we provide a heuristic framework in which a large family of kinetic equations and collective dynamics models can be regarded as the calibrated curves of a KAM function (see Section 1.4 for the finite dimensional version of this structure). We also explain how the study of the action along the calibrated curves played a crucial role in providing the intuition that led us to the emergence of phase concentration result for the Kuramoto-Sakaguchi equation. We prove this result rigorously in Theorem 3.3 from Chapter 5.

1.5.1 Emergence of phase concentration for the Kuramoto Sakaguchi equation

Collective phenomena such as aggregation, flocking, and synchronization, etc., are ubiquitous in biological, chemical, and mechanical systems in nature, e.g., the flashing of fireflies, chorusing of crickets, synchronous firing of cardiac cells, and metabolic synchrony in yeast cell suspensions. When the number of entities is sufficiently large, many of these phenomena are modeled by the Kuramoto-Sakaguchi equation:

$$\begin{cases} \partial_t f + \partial_\theta(v[f], f) = 0, \quad (\theta, \omega) \in \mathbb{T} \times \mathbb{R}, \quad t \geq 0, \\ v[f](\theta, \omega, t) = \omega - K \int_{\mathbb{T}} \sin(\theta - \theta_*)\rho(\theta_*, t)\, d\theta_*, \quad \rho(\theta, t) := \int_{\mathbb{R}} f(\theta, \omega, t)\, d\omega, \\ f(\theta, \omega, 0) = f_0(\theta, \omega), \quad \int_{\mathbb{T}} f_0 \, d\theta = g(\omega). \end{cases} \tag{1.5.34}$$

Here, $f = f(\theta, \omega)$ is the probability density function of an ensemble of oscillators in phase $\theta \in \mathbb{T}$, with a natural frequency ω at time t. Additionally, K is the positive coupling strength measuring the degree of mean-field interactions between oscillators. Lancellotti [71] rigorous derived the Kuramoto-Sakaguchi equation from the Kuramoto synchronization model by using the Neunzert's method [81].

In [20], Carrillo, Choi, Ha, Kang, and Kim showed, in the case where all the oscillators have the same frequency ω_c, (that is $f(\theta, \omega, 0) = 0$ for every $\omega \neq \omega_c$) and the diameter of the support of $\rho(\theta, 0)$ is less than π, that measured values solutions will converge to a rotating Dirac measure in the phase space (θ, ω). That is the oscillators will eventually synchronize.

More recently in [14], Benedetto, Caglioti, and Montemagno proved that, in the case of identical oscillators, any solution of (1.5.34) converges to a rotating Dirac measure in phase space, provided

$$|J_{f_0}| > 0, \tag{1.5.35}$$

where

$$J_f = \int_{\mathbb{R}} \int_{\mathbb{T}} e^{i\theta} f \, d\theta \, d\omega,$$

for any f in $\mathbb{P}(\mathbb{T} \times \mathbb{R})$.

In Chapter 5, in collaboration with Seung-Yeal Ha, Young-Heon Kim, and Jinyeong Park, in the cases of nonidentical oscillators, we show that there exists a time-dependent interval $L(t) \subset \mathbb{T}$ such that f will converge to zero uniformly outside L, provided $f_0 \in C^1(\mathbb{T} \times \mathbb{R})$, $|J_{f_0}| > 0$, and K is sufficiently large. Additionally, we quantify the diameter of L in terms of K and the support of g. By doing this, we show that the length of $L(t)$ converges to zero as $K \to \infty$.

In the identical case, that is $g(\omega) = \delta_0$, the model is a Wasserstein gradient flow of the functional

$$\mathcal{G}(f) := \frac{K}{2}(1 - |J_f|)^2.$$

Moreover, by direct computation, one finds

$$\frac{d}{dt}\mathcal{G}(f) = -KR^2 \int \sin^2(\theta - \phi) f d\theta d\omega.$$

where R and ϕ are continuous functions from $[0, \infty)$ to $\mathbb{R}$ defined by the relationship:

$$J_{f_t} = R(t)e^{i\phi(t)}.$$

Thus, it will be convinient to set

$$\mathcal{D}(f) := KR^2 \int \sin^2(\theta - \phi) f d\theta d\omega = -KR\dot{R}.$$

The functional $f \to \mathcal{D}(f)$, plays the role of $|\nabla_x E|^2$ for the gradient flow of $\mathcal{G}$ in the formalism introduced in Section 2. However, we see that the analogue of condition (1.1.5) is not satisfied by $\mathcal{G}$, i.e $\mathcal{D}(f) \to 0$ does not imply $\mathcal{G}(f) \to 0$.

In Lemma 4.1, we show that the map $t \to \mathcal{D}(f_t)$ is Lipschitz continuous. Thus, using the fact that $\mathcal{G}$ is bounded and the identity

$$\mathcal{G}(f_0) - \mathcal{G}(f_t) = \int_0^t \mathcal{D}(f_s)\, ds,$$

we show that

$$\mathcal{D}(f_t) = KR^2 \int \sin^2(\theta - \phi) f d\theta d\omega = -KR\dot{R} \to 0, \tag{1.5.36}$$

as $t \to \infty$.

Hence, we deduce that as $t \to \infty$, f should concentrate in ϕ and $-\phi$. The Wasserstein gradient of $\mathcal{G}$ at f is given by the gradient of the function $\sigma \to \sigma \cdot J_f$ in $\mathbb{S}^1$ with respect to the standard metric. For fixed f, the function $\sigma \cdot J_f$ has two critical points: a minimum at $\sigma = e^{-i\phi}$ and a maximum at $\sigma = e^{i\phi}$. Thus, if one controls the rotation of J_f it is reasonable to expect that all the mass will be concentrated in the stable equilibrium point of $\nabla\sigma \cdot J_f$ located at $\sigma = e^{i\phi}$.

Additionally, in equation (4.8) from Chapter 5, we show the estimate

$$|\dot{\phi}| \leq \sqrt{\frac{K}{R(0)}}|\dot{R}|. \tag{1.5.37}$$

Consequently, by (1.5.36), we get $|\dot{\phi}| \to 0$, as $t \to \infty$. Using this we obtain a differential inequality that shows quantitatively that solution stays away from critical points that are not global minimizers, i.e.

$$\frac{d}{dt}\int_{\cos(\theta-\phi)\leq-\delta} f^2\,d\theta \leq -C\int_{\cos(\theta-\phi)\leq-\delta} f^2\,d\theta \quad \forall t\in[T,\infty),$$

which implies

$$\int_{\cos(\theta-\phi)\leq-\delta} f_t^2\,d\theta \leq e^{-C(t-T)}\int_{\cos(\theta-\phi)\leq-\delta} f_T^2\,d\theta, \tag{1.5.38}$$

for some $C > 0$.

Such an inequality excludes concentration of mass at $e^{-i\phi}$ and yields synchronization of the oscillators in the identical case. We prove this in Theorem 3.1 of Chapter 5. This synchronization was previously obtained in [14], using different methods. However, our approach allows us to show the emergence of phase concentration in the non-identical case for a large coupling regime in Theorem 3.3 from Chapter 5, which was not known previously.

In general, (1.5.34) does not admit a gradient flow structure in the metric space $(\mathbb{P}_2(\mathbb{T} \times \mathbb{R}), W_2)$. However, in the next section, we find an action functional in the space of paths of $(\mathbb{P}_2(\mathbb{T} \times \mathbb{R}), W_2)$, where solutions to (1.5.34) can be regarded as the calibrated curves having $\mathcal{G}$ as the KAM function (see Section 1.4 for the description of these structures in finite dimensions). This KAM function is not monotonic along solutions. Nonetheless, as in the Mañe conjecture, outline in Section 1.6, periodic orbits are the expected behavior for solutions of (1.5.34). The study of the behavior of $\mathcal{G}$ along the calibrated curves plays a crucial role in the proof of Theorem 3.3. In finite dimensions, such curves converge to the projected Aubry set, which Mañe conjecture predicts to consist of equilibrium points and periodic orbits. This motivated us to study the behavior of $\mathcal{G}$ along solutions and obtain the inequality

$$\frac{d^2}{dt^2}\mathcal{G}(f(t)) \geq -4K\frac{d}{dt}\mathcal{G}(f(t)) - 2M. \tag{1.5.39}$$

Such an inequality plays the role of an entropy production estimate which was a key component in our proof. It is the key ingredient that we use in Section 6 of Chapter 5 to obtain a universal lower bound in R and show that eventually $\dot{R}$ is sufficiently small, so that the analogues of estimates (1.5.37) and (1.5.38) hold for T large enough. Then, the concentration result follows by combining these estimates with a compactness argument and the analysis of the dynamics characteristic curves of the continuity equation along solutions.

The remaining part of this Section is devoted to the heuristic framework in which (1.5.34) and a large family of kinetic equation exhibit a weak KAM structure.

1.5.2 Weak Kam structure for kinetic equations

Here, we introduce a heuristic framework in which several kinetic equations have a formal weak KAM structure. The finite dimensional version of such structure is described in Section 1.4 of the introduction. We begin by adapting the action functionals in the space of paths of measures, to phase space. In d dimensions a point in phase space is given by a couple $(x, v) \in \mathbb{R}^d \times \mathbb{R}^d$. A delta measure centered at a given point (x, v) represents a particle having position x and velocity v. Hence, if the map $t \to (x(t), v(t))$ represents the motion of a real particle, then we must have $\dot{x}(t) = v(t)$. When we consider a curve of measures $t \to f_t$ having velocity field $t \to v_t = (v_{t,x}, v_{t,v})$, i.e.

$$\partial_t f = -\mathrm{div}_x(f v_{t,x}) - \mathrm{div}_v(f v_{t,v}),$$

and minimizing our action functionals, we would like that any integral curve $t \to (x(t), v(t))$, corresponding to the flow induced by v_t, satisfies the condition $\dot{x} = v(t)$. That is, we want the curve of measures $t \to f_t$ to represent the motion of real particles in phase space. We will achieve this by using a penalization argument. We describe this argument in the general setting of the tangent space to a Riemannian manifold (M, g). This will allow us to consider a bigger family of equations and include the Cucker-Smale model, the stochastic Vicsek model, and the Kuramoto-Sakaguchi equation in Section 5.2.6.

1.5.2.1 Penalized action functionals

Let (M, g) be a Riemannian manifold and let W be a vector field in M. For each $x \in M$ we let $\mathcal{F}(x)$ be a closed, convex set in $T_x M$. Additionally, we let Λ be a positive number and μ and ν be measures in $\mathbb{P}_2(M)$. We consider the problem of minimizing

$$\int_0^\tau \int \left[\frac{1}{2}|v_t|^2 + W \cdot v_t + \frac{\Lambda}{2} \inf_{w \in \mathcal{F}(x)} \mid v_t - w \mid^2 \right] \rho_t \, dxdt, \tag{1.5.40}$$

among all measured valued maps $t \to \rho_t$ and vector fields $t \to v_t$ from $[0, \tau]$ to $\mathbb{P}_2(M)$ such that $\rho_0 \, dx = \mu$, $\rho_\tau \, dx = \nu$, and

$$\frac{d}{dt} \int \zeta \rho \, dx = \int \nabla \zeta v_t \rho_t \, dx \quad \forall \zeta \in C^\infty(M) \quad \text{and} \quad \forall t \in [0, \tau]. \tag{1.5.41}$$

Here, all the norms, inner products, gradients and divergences will be taken with respect to g. Additionally, we assume the assignation $x \to \mathcal{F}(x)$, is well behaved. By this we mean that the map $P : TM \to TM$, given by

$$P(x, v) = \left(x, \text{argmin}_{w \in \mathcal{F}(x)} \frac{1}{2} \mid v - w \mid^2 \right),$$

is differentiable.

1.5.2.2 Non-coercive Hamiltonians

To understand minimizers of (1.5.40) when $\Lambda \to \infty$, it will be convenient to compute the Hamiltonian associated with the Lagrangian

$$L(x, v) = \frac{1}{2}|v|^2 + W \cdot v + \frac{\Lambda}{2} \inf_{w \in \mathcal{F}(x)} \mid v - w \mid^2 .$$

For this purpose, in Section G.2 from Chapter 5, we show that the Hamiltonian H associated with L is given by

$$
\begin{aligned}
H(x,p) = \sup_{w\in\mathcal{F}(x)} p\frac{p-W+\Lambda w}{1+\Lambda} - \frac{1}{2}\left|\frac{p-W+\Lambda w}{1+\Lambda}\right|^2 - W\cdot\frac{p-W+\Lambda w}{1+\Lambda} \\
- \left|\frac{p-W+\Lambda w}{1+\Lambda} - w\right|^2.
\end{aligned}
\tag{1.5.42}
$$

Additionally, we show that when we let $\Lambda \to \infty$ in (1.5.42), we obtain, $H \to \mathcal{H}$, where

$$
\mathcal{H}(x,p) = pP(x,p-W) - W\cdot P(x,p-W) - \frac{1}{2} \mid P(x,p-W) \mid^2 .
$$

Then,

$$
\begin{aligned}
\mathcal{L}(x,v) &= \sup_p pv - pP(x,p-W) - P(x,p-W)\cdot W + \frac{1}{2} \mid P(x,p-W) \mid^2 \\
&= \begin{cases} \frac{1}{2} \mid P(x,p-W) \mid^2 - P(x,p-W)\cdot W & \text{if} \quad P(x,p-W) = v, \\ \infty & \text{otherwise.} \end{cases}
\end{aligned}
$$

Consequently, if $\mathcal{L}(x,v)$ is finite and v and p are conjugates via the Legendre transform (see Section 1.4), then we must have

$$
v = P(x,p-W).
$$

Here, we are abusing notation to regard the covector p as a vector. To do this we are using the canonical isomorphism between TM^* and TM induced by g. From now on, we will adopt this convention.

1.5.2.3 Optimal velocity fields as an infinite dimensional Legendre transform

Suppose the curve $t \to \rho_t$ with velocity field $t \to v_t$, is a minimizer of (1.5.40) and fix $t \geq 0$. As in Section 2.2 of the introduction, we note that if $t \to v_t, \rho_t$ satisfies constraint (1.5.41), so does $v_t + sA$, provided $\operatorname{div}(A\rho_t) = 0$. Consequently, by minimality we must have

$$0 = \frac{d}{ds}\bigg|_{s=0} \int \Big[L(x, v_t + sA)\Big]\rho_t \, dx = \int \nabla_v L(x, v_t) A \, dx.$$

Since this hold for any A with the property that $\operatorname{div}(A\rho_t) = 0$, by the Helmholtz-Hodge Theorem there exists $\nabla\varphi_t$ with the property that

$$\nabla_v L(x, v_t) = \nabla\varphi_t. \tag{1.5.43}$$

Using Legendre duality we conclude

$$\nabla_p H(x, \nabla\varphi_t) = v_t.$$

As $\Lambda \to \infty$, the discussion of the previous section yields

$$v_t = P(x, \nabla\varphi_t - W).$$

In the case where $\mathcal{F}(x)$ is a subspace, this is just the orthogonal projection of $\nabla\varphi_t - W$ into $\mathcal{F}(x)$.

Relationship (1.5.43) can be regarded as an infinite dimensional Legendre transform. Indeed, in finite dimensions, if v_p is the Legendre conjugate of p then,

$$v_p = \operatorname{argmax}_v \, pv - L(x, v).$$

On the other hand, when we use the formal Riemannian structure of Felix Otto (see Section 2), with

$$T_\rho \mathbb{P}_2(M) = \overline{\left\{ \nabla\varphi \;:\; \int |\nabla\varphi|^2 \rho \, dx < \infty \right\}}^{L^2(\rho)}$$

and we consider the Legendre conjugate of a covector $\nabla\varphi$, we maximize

$$\omega \to \int \nabla\varphi \nabla\omega \rho \, dx - \int L(x, \nabla\omega)\rho \, dx.$$

If ω is a maximizer of the above map, then we must have that for any function α,

$$\begin{aligned} 0 &= \frac{d}{ds} \int \nabla\varphi \nabla(\omega + s\alpha)\rho \, dx - \int L(x, \nabla(\omega + s\alpha))\rho \, dx \\ &= \int \nabla\varphi \nabla\alpha \rho \, dx - \int \nabla_v L(x, \nabla\omega) \nabla\alpha \rho \, dx. \end{aligned}$$

Since α was arbitrary, we must have

$$\nabla\varphi = \nabla_v L(x, \nabla\omega),$$

which is equivalent to (1.5.43).

1.5.2.4 Extremals of the action

In this section, we show a heuristic argument that characterizes minimizers of (1.5.40) and their limiting properties as $\Lambda \to \infty$. The main idea is to generalize the techniques of Section 3 of the introduction to compute the first variation of paths in $\mathbb{P}_2(M)$ minimizing (1.5.40). Let $t \to \rho_t$, be a minimizer of

(1.5.40), and let $t \to \varphi_t$ be the potential generating the corresponding optimal velocity field indexed in $[0, \tau]$ as described in the previous section so that

$$\frac{d}{dt}\int \zeta\rho_t\, dx = \int \nabla\zeta\nabla_p H(x, \nabla\varphi_t)\rho_t\, dx \quad \forall\zeta \in C^\infty(M) \quad \text{and} \quad \forall t \in [0, \tau],$$

(see Section 5.2.3).

Here, we will we show that $t \to \varphi_t$, satisfies

$$\partial_t\varphi_t + H(x, \nabla\varphi_t) = 0 \quad \text{in} \quad [0, \tau], \tag{1.5.44}$$

where H is the Hamiltonian corresponding to the Lagrangian

$$L(x, v) = \frac{1}{2}|v|^2 + v \cdot W + \frac{\Lambda}{2} \inf_{w\in\mathcal{F}(x)} \mid v - w \mid^2 .$$

Hence, by the discussion of Sections 5.2.2 and 5.2.3, we get that when $\Lambda \to \infty$, minimizers satisfy

$$\partial_t\varphi + \nabla\varphi P(x, \nabla\varphi - W) - W \cdot P(x, \nabla\varphi - W) - \frac{1}{2} \mid P(x, \nabla\varphi - W) \mid^2 = 0,$$

and

$$\frac{d}{dt}\int \zeta\rho_t\, dx = \int \nabla\zeta P(x, \nabla\varphi - W)\rho_t\, dx \quad \forall\zeta \in C^\infty(M) \quad \text{and a.e } t \quad \text{in} \quad [0, \tau].$$

In order to see this, we perturb such minimizers by building a vector field along the minimizing path, (as we did in Section 3 of the introduction). That is, for each t in $[0, \tau]$ we consider a potential ω_t and the vector field $\nabla\omega_t$ induced by it. We require ω_t to be identically 0 in the complement of a compact subset of $(0,\tau)$, so that the variation fixes the end points of the optimal path.

Indeed, for each s, we let $t \to \rho_{t,s}$ and $t \to \nabla_p H(x, \nabla\varphi_{t,s})$ satisfy

$$\frac{d}{dt}\int \zeta\rho_{t,s}dx = \int \nabla\zeta\nabla_p H(x,\nabla\varphi_{t,s})\rho_{t,s}\, dx \quad \forall\zeta \in C^\infty(M),$$

and

$$\frac{d}{ds}\bigg|_{s=0}\int \zeta\rho_{t,s}\, dx = \int \nabla\zeta\nabla\omega_t\rho_{t,s}\, dx \quad \forall\zeta \in C^\infty(M).$$

Then, by minimality of the path $t \to \rho_t$, when $s = 0$, we must have

$$0 = \frac{d}{ds}\int_0^\tau \int L(x, \nabla_p H(x, \nabla\varphi_{t,s})\rho_{t,s}\, dx.$$

In Section G.4 of Chapter 5, we show that this condition is equivalent to the identity,

$$0 = -\int_0^\tau \int \left[\nabla\omega_t\nabla[\partial_t\varphi_t + H(x,\nabla\varphi_t)]\rho_t\, dx\right] dt.$$

Since $t \to \omega_t$ was arbitrary in compact subsets of $(0,\tau)$, (1.5.44) follows.

1.5.2.5 Minimizing movement scheme of the entropy

In this section, we consider an infinite dimensional variant of the method described in Section 1.5 of the introduction to build calibrated curves of a KAM function starting at a point x_0. By doing this, in the next section, we will be able to regard a large family of kinetic equations as calibrated curves of KAM functions.

Let ρ_0 be a probability density in $\mathbb{P}_2(M)$. We provide a heuristic argument to characterize minimizers of

$$\{\rho_t\}_{t\in[0,\tau]} \to \int_0^\tau \frac{1}{2}\int \left[|v_t|^2 + W\cdot v_t + \frac{\Lambda}{2}\inf_{\omega\in\mathcal{F}(x)}|v_t-\omega|^2\right]\rho_t\,dxdt + \int \rho_\tau\log\rho_\tau\,dx, \tag{1.5.45}$$

as well as their properties when $\Lambda\to\infty$. Here, the vector fields $t\to v_t$ satisfy constraint (1.5.41) and ρ_τ is a free parameter in the minimization.

In Section 5.2.3 we saw that any optimal velocity field has, for each t in $[0,\tau]$, a function φ_t such that

$$v_t = \nabla_p H(x,\nabla\varphi_t).$$

In Section 5.2.4 we found that any potential $t\to\varphi_t$ corresponding to an optimal path satisfies

$$\partial_t\varphi + H(x,\nabla\varphi_t) = 0,$$

and

$$\frac{d}{dt}\int\zeta\rho_t dx = \int\nabla\zeta\nabla_p H(x,\nabla\varphi_t)\rho_t\,dx \quad \forall\zeta\in C^\infty(M) \quad \text{and} \quad \forall t\in[0,\tau].$$

In Section G.5 of Chapter 5 we will show that minimizers of (1.5.45) satisfy

$$\varphi_\tau = -\log\rho_\tau - 1 + c,$$

for some constant c.

Hence, by the discussion of Section 5.2.3, using the minimizing movement scheme and letting $\Lambda \to \infty$, one builds solutions to

$$\partial_t \rho = -\mathrm{div}(\rho P(x, -\nabla \log \rho - W)).$$

1.5.2.6 Examples

The Cucker-Smale model. We let $(M, g) = \mathbb{R}^d \times \mathbb{R}^d$, with the standard product metric. Additionally, we set $W = 0$, $\mathcal{F}(x, v) = \{v\} \times \mathbb{R}^d \subset T_{(x,v)}(\mathbb{R}^d \times \mathbb{R}^d)$. Then, by Section 5.2.3, when $\Lambda \to \infty$, we have

$$P\big((x, v), \nabla\varphi\big) = (v, \nabla_v \varphi).$$

Hence, by Section 5.2.4, optimal paths satisfy

$$\begin{aligned} \frac{d}{dt} \int \zeta f_t d\omega d\sigma &= \int \nabla\zeta P\big((x, v), \nabla\varphi_t\big) f_t \, d\omega d\sigma \\ &= \int \big[v \nabla_x \zeta + \nabla_v \zeta \nabla_v \varphi_t\big] f_t \, d\omega d\sigma, \end{aligned}$$

for any $\zeta \in C^\infty(M)$ and

$$\begin{aligned} 0 &= \partial_t \varphi + \nabla\varphi P(x, \nabla\varphi) - \frac{1}{2} \mid P(x, \nabla\varphi) \mid^2 \\ &= \partial_t \varphi + v \cdot \nabla_x \varphi + \frac{1}{2} |\nabla_v \varphi|^2 - \frac{1}{2} |v|^2. \end{aligned}$$

Then, proceeding as in Section 5.2.5 when one minimizes

$$\begin{aligned} \{f_t\}_{t \in [0,\tau)} \to \frac{1}{4} \int \psi(|x - x^*|) |v - v^*|^2 f_\tau(x, v) f_\tau(x^*, v^*) \, dx dv dx^* dv^* \\ + \int_0^\tau \left(\int L(x, v_t) f_t \, d\omega d\sigma \right) dt, \end{aligned}$$

with f_0 fixed, one finds the optimality condition

$$\varphi_\tau(x, v) = -\frac{1}{2} \int \psi(|x - x^*|) |v - v^*|^2 f_\tau(x^*, v^*) \, dx^* dv^* + c,$$

for some constant c.

Consequently, when one follows the minimal movement scheme and lets $\Lambda \to \infty$, one should obtain solutions to

$$\begin{aligned}\partial_t f &= -\mathrm{div}\Big(fP\big(x, \nabla \frac{1}{2}(\int \psi(|x-x^*|)|v-v^*|^2 dx^* dv^*)\big)\Big)\\&= -\mathrm{div}_x\big(vf\big) - K\mathrm{div}_v\big(f\nabla_v \frac{1}{2}\int \psi(|x-x^*|)|v-v^*|^2 dx^* dv^*\big)\\&= -\mathrm{div}_x\big(vf\big) - K\mathrm{div}_v\big(fF_a(f)\big),\end{aligned}$$

where

$$F_a(f) = \int \psi(|x-x^*|)(v^*-v)f \, dx^* dv^*.$$

Stochastic Vicsek model. We let $(M,g) = \mathbb{R}^d \times \mathbb{S}^{d-1}$ with the standard product metric. Additionally, we set $W = 0$ and $\mathcal{F}(x,\omega) = \{\omega\} \times \{T_\omega \mathbb{S}^{d-1}\} \subset T_{(x,\omega)}(\mathbb{R}^d \times \mathbb{S}^{d-1})$. Then, by Section 5.2.2, when $\Lambda \to \infty$ we have

$$P(x, \nabla\varphi) = (\omega, \nabla_\omega \varphi).$$

Here, ∇_ω denotes the gradient in $\mathbb{S}^{d-1}$ with the standard metric.

Hence, by Section 5.2.4, optimal paths satisfy

$$\begin{aligned}\frac{d}{dt}\int \zeta f_t \, dx d\omega = \int \nabla\zeta P(x, \nabla\varphi_t) f_t \, dx d\omega = \int \nabla_x \zeta \omega f_t \, dx dv\\ + \int \nabla_\omega \zeta \nabla_\omega \varphi_t f_t \, dx dv\end{aligned}$$

$\forall \zeta \in C^\infty(M)$,

and

$$0 = \partial_t\varphi + \nabla\varphi P(x,\nabla\varphi) - \frac{1}{2} \mid P(x,\nabla\varphi) \mid^2 = \partial_t\varphi + \omega\cdot\nabla_x\varphi + \frac{1}{2} \mid \nabla_\omega\varphi \mid^2 - \frac{1}{2}.$$

Then, proceeding as in Section 5.2.5 when one minimizes

$$\{f_t\}_{t\in[0,\tau)} \to \int \Big(f_\tau \log f_\tau - \mid J_{f_\tau}(x) \mid \Big) dxdv + \int_0^\tau \Big(\int L(x,v_t) f_t dv d\omega \Big) dt.$$

with f_0 fixed, one finds the optimality condition

$$\varphi_\tau = -\log \frac{f_\tau}{e^{\omega\cdot\Omega_{f_\tau}}} + c,$$

for some constant c. Here

$$J_f(x) = \int \omega f(x,\omega) d\omega,$$

and

$$\Omega_f(x) = \frac{J_f(x)}{\mid J_f(x) \mid}.$$

Consequently, when one follows the minimal movement scheme and lets $\Lambda \to \infty$, one should obtain solutions to

$$\begin{aligned}\partial_t f &= -\mathrm{div}(fP(x, -\nabla \log \frac{f_\tau}{e^{\omega\cdot\Omega_{f_\tau}}})) \\ &= -\mathrm{div}_x(\omega f) + \mathrm{div}_\omega\big(f\nabla \log \frac{f}{e^{\omega\cdot\Omega_f}}\big) \\ &= -\mathrm{div}_x(\omega f) + \Delta_\omega f - \mathrm{div}_\omega(f\nabla_\omega\omega\cdot\Omega_f).\end{aligned}$$

Kuramoto Synchronization model. We let $(M,g) = \mathbb{R}\times\mathbb{S}^1$, with the standard product metric. Additionally, we set $W(\omega,\sigma) = (0, -\omega\sigma^\perp) \in$

$T_{(\omega,\sigma)}(\mathbb{R}\times\mathbb{S}^1)$, $\mathcal{F}(\omega,\sigma) = \{0\}\times\{T_\sigma\mathbb{S}^1\} \subset T_{(\omega,\sigma)}(\mathbb{R}\times\mathbb{S}^1)$. Then, by Section 5.2.3, when $\Lambda\to\infty$, we have

$$P(x,\nabla\varphi) = (0,\nabla_\sigma\varphi).$$

Here, ∇_σ denotes the gradient in $\mathbb{S}^1$ with the standard metric.

Hence, by Section 5.2.4, optimal paths satisfy

$$\begin{aligned}\frac{d}{dt}\int \zeta f_t d\omega d\sigma &= \int \nabla\zeta P\big(x,\nabla\varphi_t + (0,\omega\sigma^\perp)\big) f_t\, d\omega d\sigma\\ &= \int \nabla_\sigma\zeta\cdot\big[\omega\sigma^\perp + \nabla_\sigma\varphi\big] f_t\, d\omega d\sigma,\end{aligned}$$

and

$$\begin{aligned}0 &= \partial_t\varphi + \nabla\varphi P(x,\nabla\varphi + (0,\omega\sigma^\perp)) + (0,\omega\sigma^\perp)\cdot P(x,\nabla\varphi+(0,\omega\sigma^\perp))\\ &\quad - \frac{1}{2}\mid P(x,\nabla\varphi+(0,\omega\sigma^\perp))\mid^2\\ &= \partial_t\varphi + \frac{1}{2}|\nabla_\sigma\varphi|^2 + \omega\sigma^\perp\cdot\nabla_\sigma\varphi + \frac{1}{2}\omega^2,\end{aligned}$$

for any $\zeta\in C^\infty(M)$ and t in $[0,\infty)$.

Then, proceeding as in Section 5.2.5 when one minimizes

$$\{f_t\}_{t\in[0,\tau)} \to \frac{K}{2}\Big(1 - |J_{f_\tau}|^2\Big) + \int_0^\tau \left(\int L(x,v_t) f_t\, d\omega d\sigma\right) dt,$$

with f_0 fixed, one finds the optimality condition

$$\varphi_\tau(\omega,\sigma) = K\sigma\cdot J_{f_\tau} + c,$$

for some constant c.

Here,

$$J_f(x) = \int \sigma f(\sigma, \omega)\, d\sigma d\omega,$$

Consequently, when one follows the minimal movement scheme and lets $\Lambda \to \infty$, one should obtain solutions to

$$\begin{aligned}\partial_t f &= -\mathrm{div}(fP(x, K\nabla\sigma \cdot J_{f_\tau} + (0, \omega\sigma^\perp))) \\ &= -K\mathrm{div}_\sigma\big(f\nabla_\sigma\sigma \cdot J_{f_\tau}\big) - K\mathrm{div}_\sigma\big(f\omega\sigma^\perp\big),\end{aligned}$$

which is the Kuramoto Sakaguchi equation in $\mathbb{R} \times \mathbb{S}^1$.

1.6 Formal weak KAM structure for reaction-diffusion equations with boundary conditions

In this section, we introduce a heuristic framework in which a large family of reaction-diffusion equations with boundary conditions regarded as the calibrated curves of a KAM function. This family includes the Fokker-Planck equation with positive Dirichlet and Robin boundary conditions, population growth models, and the Fisher-KPP equation. We also explain how this structure allowed us to adapt the transportation cost found by Alessio Figalli and Nicola Gigli in [43], to build solutions to reaction-diffusion equations with positive Dirichlet boundary condition by using the minimizing movement scheme. We prove this result rigorously in Theorem 4.1 from Chapter 2.

1.6.1 A new family of transportation costs with applications to reaction-diffusion equations with boundary conditions

In general, the gradient flow interpretation for evolution equations in a suitable space X provides a powerful variational method, developed by De

Giorgi, to prove the existence of solutions. Given a time step $\tau > 0$ and an initial configuration $x_0 \in X$, one builds an approximate solution by iteratively minimizing.

$$x \to \frac{d^2(x_n, x)}{2\tau} + E(x) = E[x|x_n],$$

where x_n is a minimum for $E[x|x_{n-1}]$, and d is a distance defined in X.

This method has a wide range of applications, from the work of Almgren, Taylor, and Wang on mean curvature flow [4] to kinetic equations [19].

Whenever X is the space of probability measures on an open domain, and d is the Wasserstein distance, the minimizing movement scheme will always produce solutions to parabolic equations with Neumann boundary conditions. This interpretation allows one to prove entropy estimates and functional inequalities (see [84, 94] for more details on this area, which is very active and in constant evolution).

More recently, Figalli and Gigli [43] introduced a distance among positive measures in an open domain Ω. Such a distance allows one to use this scheme to build solutions to the problem

$$\begin{cases} \partial_t \rho = \operatorname{div}\Big(\nabla\rho + \rho\nabla V\Big) \quad \text{in } \Omega, \\ \rho = e^{-V} \quad \text{on} \quad \partial\Omega. \end{cases} \tag{1.6.46}$$

Their work is a pioneering effort in the project of extending the methods developed by Otto to study evolution equations with Dirichlet boundary conditions.

They focus on the case when the energy is the relative entropy to illustrate the main ideas. Indeed, when E is the relative entropy with respect to e^{-V}, their approach yields solutions to (1.6.46). As in the Wasserstein case if one uses other functionals in their construction one obtains solutions of a large family of equations subject to Dirichlet boundary conditions. Examples of these include porous medium equations and nonlocal interaction equations.

When one uses this approach one would like to be able to choose the boundary condition. However, as in (1.6.46), when one uses the distance introduced by Figalli and Gigli [43] the boundary condition and the equation are always coupled.

In Chapter 2, we introduce transportation costs between positive measures on domains that allow the equation and the boundary condition to be detached. More precisely, this chapter shows that when the distance is replaced by one of these costs in the minimizing movement scheme, one obtains solutions to

$$\begin{cases} \partial_t \rho = \operatorname{div}\Big(\nabla\rho + \rho\nabla V\Big) - h(\rho) \quad \text{in} \quad \Omega, \\ \rho = \rho_D \quad \text{on} \quad \partial\Omega, \end{cases} \tag{1.6.47}$$

where ρ_D is strictly positive, and h belongs to a large family of reaction terms. Note that, because of the drift term, it is not possible to choose an energy functional to obtain such an equation as gradient flow in $L^2(\Omega)$. In contrast to other transportation distances among positive measures [72, 80], the costs in

Chapter 2 allow the diffusion term and the reaction term to be set independently.

To decouple the boundary condition and the equation we develop a heuristic framework that extends Otto's formalism to the case when the measures exchange mass with the boundary. In fact, the class of costs we introduced is not associated with distances. We overcame this difficulty by using the formal manifold structure of the space of paths of measures that interact with the boundary. These heuristic arguments (discussed in Section 2 of [79]) allow one to consider general reaction terms and boundary conditions. We found these heuristics by combining the works of Otto with the work of Milnor [78] on the formal Riemannian structure in the space of paths, the work of Rossi and Piccoli on the generalization of the Benamou Brenier formula for positive measures [11], and a paper by Figalli, Gangbo, and Yolcu on the minimizing movement scheme with Lagrangian cost [42].

The remaning parts of this section are devoted to this heuristic framework.

1.6.2 Weak KAM structure for reaction-diffusion equations with boundary conditions

Here, we introduce a heuristic framework in which several reaction-diffusion and parabolic equations have a formal weak KAM structure. The finite dimensional version this structure is described in Section 1.4 of the introduction. We begin by using an action functional in the space of paths of positive measures in bounded domains that interact with the boundary to define a transportation cost. For this purpose we denote by $\mathcal{M}(\Omega)$ the space of positive measures in Ω and we consider the following problem:

Problem 2.1 (A variant of the transportation problem). *Given μ and ν in $\mathcal{M}(\Omega)$ we consider the problem of minimizing*

$$\tilde{C}_\tau(v_t, h_{t,}, \overline{h_t}) = \int_0^\tau \left[\frac{1}{2}\int_\Omega |v_t|^2 \rho_t \, dx + \int_\Omega e(h_t) m(\rho_t)\, dx + \int_{\partial\Omega} \overline{e}(\overline{h}_t)\, d\mathcal{H}^{d-1}\right] dt,$$

among all positive measured valued maps from $[0,\tau]$ to $\mathcal{M}(\Omega)$, satisfying $\rho_0 \, dx = \mu$ and $\rho_\tau \, dx = \nu$. Here, the measures ρ_t and the triplets $(v_t, h_{t,}, \overline{h_t})$ are indexed by t in $[0,\tau]$. We require them to satisfy the constraint

$$\frac{d}{dt}\int_\Omega \zeta\rho_t \, dx = \int_\Omega \langle \nabla\zeta, v_t\rangle \rho_t \, dx - \int_\Omega \zeta h_t m(\rho)\, dx - \int_{\partial\Omega} \zeta \overline{h}_t \, d\mathcal{H}^{d-1},$$
$$\forall t \in [0,\tau] \text{ and } \forall \zeta \in C^\infty(\overline{\Omega}).$$

This provides a transportation cost between μ and ν given by

$$Wb_2^{e,m,\bar{e},\tau}(\mu,\nu) := \inf \tilde{C}_\tau(v_t, h_{t,}, \overline{h_t}).$$

Henceforth, a path is defined as a measured valued map from $[0,\tau]$ to $\mathcal{M}(\Omega)$. We apply the minimizing movement scheme to this cost: Given an initial measure ρ_0, we build a family of curves $t \to \rho^\tau(t)$, indexed by $\tau > 0$, iterating the minimization of the map

$$\rho \to \int_\Omega [\rho\log\rho - \rho + V(x)\rho + 1]\, dx + Wb_2^{e,m,\bar{e},\tau}(\rho_n^\tau, \rho) = \tilde{E}^\tau[\rho|\rho_n^\tau],$$

where ρ_n^τ is a minimum of $\tilde{E}^\tau[\rho|\rho_{n-1}^\tau]$, in $\mathcal{M}(\Omega)$. We define the discrete solutions by

$$\rho^\tau(t) := \rho^\tau_{[t/\tau]}.$$

Then, as $\tau \downarrow 0$, we extract a subsequence converging to a weak solution of the problem:

$$\begin{cases} \partial_t \rho = \operatorname{div}\Big(\nabla\rho + \rho\nabla V\Big) - [e_x']^{-1}(\log\rho + V)m(\rho), \quad in \quad \Omega, \\ -\langle \nabla\rho - \nabla V\rho, \nu\rangle = [\overline{e}_x']^{-1}(\log\rho + V) \quad in \quad \partial\Omega, \\ \rho(0) = \rho_0. \end{cases}$$

In particular, when we set

$$e(h,x) = \begin{cases} \int_{[F_x']^{-1}(0)}^{[F_x']^{-1}(h)} (\log r + V) F_x''(r)\, dr, \quad \text{if} \quad [F_x']^{-1}(h) \geq 0, \\ +\infty, \quad \text{otherwise}, \end{cases}$$

and

$$\overline{e}(\overline{h},x) = \begin{cases} g_R\Big(l(\overline{h})\log l(\overline{h}) + (V-1)l(\overline{h}) + 1\Big), \quad \text{if} \quad l(\overline{h}) \geq 0, \\ +\infty, \quad \text{otherwise}, \end{cases}$$

we obtain a weak solution to the problem:

$$\begin{cases} \partial_t \rho = \operatorname{div}\Big(\nabla\rho(t) + \rho(t)\nabla V\Big) - F_x'(\rho)m(\rho) \quad \text{in} \quad \Omega, \\ -\langle \nabla\rho - \nabla V\rho, \nu\rangle = g_R(\rho - \rho_R) \quad \text{on} \quad \partial\Omega. \end{cases} \tag{1.6.48}$$

Here,

$$l(r) = \frac{r}{g_R} + \rho_R.$$

Also, we will show that when we set

$$e(h,x) = \begin{cases} \int_{[F_x']^{-1}(0)}^{[F_x']^{-1}(h)} (\log r + V) F_x''(r)\, dr, \quad \text{if} \quad [F_x']^{-1}(h) \geq 0, \\ +\infty, \quad \text{otherwise}, \end{cases}$$

and

$$\overline{e}(\overline{h}, x) = (\log \rho_D + V)\overline{h},$$

we obtain a weak solution to

$$\begin{cases} \partial_t \rho = \operatorname{div}\Big(\nabla \rho(t) + \rho(t)\nabla V\Big) - F_x'(\rho)m(\rho) \quad \text{in} \quad \Omega, \\ \rho = \rho_D \quad \text{on} \quad \partial\Omega. \end{cases} \tag{1.6.49}$$

When $F_x'(r) = (r-1)$, $m(r) = r$, and $V = 0$, one obtain solution of to the Fisher-KPP equation.

The heuristic is presented as follows. Section 6.2.1 characterizes optimal triplets in terms of potentials. Section 6.2.2 describes a characterization of minimal paths in terms of an equation for the potentials. Section 6.2.3 describes how the equation for minimal paths can be used to perform the minimizing movement scheme. Section 6.2.4 describes how to match the cost with the boundary conditions. Section 6.2.5 describes how to match the cost with the reaction term. Finally, Section 6.2.6 show discusses a heuristic computation in which the action between the calibrated curves of the entropy decreases exponentially in time. This is another infinite dimensional analogue of the structure introduced in Section 1.4.

1.6.2.1 Optimal triplets

In Section 2.1 of Chapter 2 we show a heuristic argument that characterizes minimizing triplets for Problem 2.1. For such triplets, there exist functions φ_t indexed in $[0,\tau]$, such that:

(a) $\nabla\varphi_t = v_t$.

(b) $\varphi_t = -\overline{e}'(\overline{h}_t)$ on $\partial\Omega$ and $\overline{h_t} = \langle \rho\nabla\varphi_t, \nu\rangle$.

(c) $\varphi_t = -e'(h_t) \quad \text{in} \quad \Omega$.

As in Section 2.2 of the introduction, such a potential is found by performing variations of the triplets that preserve constraint (**??**)

1.6.2.2 Optimal paths

In Section 2.2 of Chapter 2, we show a heuristic argument that characterizes minimizers of Problem 2.1.

Let ρ_t, indexed in $[0,\tau]$, be a minimizer of Problem 2.1. Also, for each t in $[0,\tau]$, let φ_t be the potential generating the corresponding optimal triplet: $\big(\nabla\varphi_t, [e']^{-1}(-\varphi_t), [\overline{e}']^{-1}(-\varphi_t)\big)$. Then, we show that

$$\partial_t\varphi + \frac{1}{2}|\nabla\varphi_t|^2 - \big[\varphi_t[e']^{-1}(-\varphi_t) + e([e']^{-1}(-\varphi_t))\big]m'(\rho_t) = 0, \qquad (1.6.50)$$

and

$$\frac{d}{dt}\int_\Omega \zeta\rho_t\,dx = \int_\Omega \langle\nabla\zeta, \nabla\varphi_t\rangle\rho_t\,dx - \int_\Omega \zeta[e']^{-1}(-\varphi_t)m(\rho_t)\,dx \\ - \int_{\partial\Omega} \zeta[\overline{e}']^{-1}(-\varphi_t)\,d\mathcal{H}^{d-1},$$

for every ζ in $C_c^\infty(\overline{\Omega})$.

As in Section 5.2.4 of the introduction we obtain this condition by using a vector field along the curve of measures to induced a variation with fixed end points (see Section 3 of the introduction to the differential structure for path space).

1.6.2.3 The minimizing movement scheme.

Given $\rho_0 \in \mathcal{M}(\Omega)$ and $\tau > 0$, in Section 2.3 of Chapter 2 we provide arguments to characterize the minimizers of

$$\{\rho_t\}_{t\in[0,\tau]} \to \int_0^\tau \left(\frac{1}{2}\int_\Omega |v_t|^2\rho_t\,dx + \int_\Omega e(h_t)m(\rho_t)\,dx + \int_{\partial\Omega} \overline{e}(\overline{h}_t)\,d\mathcal{H}^{d-1}\right)dt \\ + \int_\Omega \left[\rho_\tau \log\rho_\tau + \big(V-1\big)\rho_\tau + 1\right]dx. \tag{1.6.51}$$

Here, the triplets $(v_t, h_t, \overline{h}_t)$ satisfy (**??**). Also, ρ_0 is fixed and $\rho_\tau = \rho$.

In Section 6.2.1 we saw that for minimizing triplets we have for each $t \in [0,\tau]$ a function φ_t such that

$$(v_t, h_t, \overline{h}_t) = \big(\nabla\varphi_t, [e']^{-1}(-\varphi_t), [\overline{e}']^{-1}(-\varphi_t)\big).$$

In Section 6.2.2 we found that optimal paths satisfy

$$\partial_t\varphi_t + \frac{1}{2}|\nabla\varphi_t|^2 - \big[\varphi_t[e']^{-1}(-\varphi_t) + e([e']^{-1}(-\varphi_t))\big]m'(\rho_t) = 0.$$

In Section 2.3 of Chapter 2 we will show that minimizers of (1.6.51) must satisfy

$$\varphi_\tau = -\log \rho_\tau - V.$$

In order to see this, we suppose that we have a minimizer ρ_τ and a path $t \to \rho_t$, with corresponding triplets $t \to \big(\nabla\varphi_t, [e']^{-1}(-\varphi_t), [\overline{e}']^{-1}(-\varphi_t)\big)$.

Proceeding as in Section 1.5 of the introduction, we obtain this condition by using a vector field along the curve of measures to induced a variation with that fixes ρ_0 but lets ρ_τ be free (see Section 3 of the introduction to the differential structure for path space and Section 2.5 of Chapter 2 for the computation).

1.6.2.4 The Boundary Conditions

In Section 6.2.3 we showed that for minimizers of (1.6.51) we have that $\varphi_\tau = -\log \rho_\tau - V$. In Section 6.2 we showed that for optimal triplets

$$-\varphi = \overline{e}'(\overline{h}) \quad \text{on} \quad \partial\Omega.$$

Hence, if we set $\overline{e}(\overline{h}) = \Psi\, \overline{h}$, we obtain the boundary condition

$$\rho_\tau = e^{\Psi - V} \quad \text{on} \quad \partial\Omega.$$

This concludes the analysis for the boundary condition for (1.6.49).
In order to derive the boundary condition for (1.6.48), we proceed as follows: In Section 2.1 of Chapter 2, we show that for minimizers of Problem 2. we

must have

$$\overline{h}_t = \langle \rho_t \nabla \varphi_t, \nu \rangle.$$

Hence, we expect the limit of the minimizing movement scheme to satisfy the relation

$$-\langle \nabla \rho, \nu \rangle - \langle \rho \nabla V, \nu \rangle = [\overline{e}']^{-1}(\log \rho + V).$$

Our goal is to obtain the boundary condition

$$-\langle \nabla \rho, \nu \rangle - \langle \rho \nabla V, \nu \rangle = g_R\big(\rho - \rho_R\big).$$

For this purpose, we would need

$$[\overline{e}']^{-1}(\log \rho + V) = g_R\big(\rho - \rho_R\big).$$

Thus,

$$V + \log \rho = [\overline{e}']\Big(g_R\big(\rho - \rho_R\big)\Big).$$

Hence, if we set

$$l(r) = \frac{r}{g_R} + \rho_R,$$

we obtain

$$[\overline{e}'](l^{-1}(\rho)) = \log \rho + V.$$

Then, it follows that

$$[\overline{e}'](l^{-1}(\rho))[l^{-1}(\rho)]' = g_R(\log \rho + V).$$

Integrating, we get

$$\overline{e}(l^{-1}(\rho)) = \int_0^\rho g_R\big(\log r + V\big) dr + C.$$

Here, C is a constant that will be chosen later. This implies

$$\overline{e}(\rho) = g_R \int_{l(0)}^{l(\rho)} \big(\log r + V\big) dr + C.$$

Thus, it suffices to set

$$\overline{e}(\rho) = g_R \bigg(l(\rho) \log l(\rho) + \bigg(V - 1\bigg) l(\rho) + 1 \bigg),$$

where C has been chosen appropriately.

1.6.2.5 The reaction term

In Section 6.2.1 we showed that optimal triplets satisfy

$$e'(h_\tau) = -\varphi_\tau.$$

In Section 6.2.3 we showed that minimizers of (1.6.51) satisfy

$$\varphi_\tau = -\log \rho_\tau - V.$$

Thus, in order to obtain

$$h = F'(\rho),$$

we set

$$e'(F'(\rho)) = \log \rho + V.$$

This implies

$$e'(F'(\rho)) F''(\rho) = \big(\log \rho + V\big) F''(\rho).$$

Integrating we obtain

$$e(F'(\rho)) = \int_0^\rho \big(\log r + V\big) F''(r)\, dr + C,$$

for some constant C. Thus, it suffices to set

$$e(\rho) = \int_{[F']^{-1}(0)}^{[F']^{-1}(\rho)} \big(\log r + V\big) F''(r)\, dr,$$

where C has been chosen appropriately.

1.6.2.6 Stability estimates

Here we provide an example in which the minimizing movement scheme yields a contractible flow. By this we mean that the cost between any two solutions of the flow decreases exponentially in time. Here, $\Omega = \mathbb{R}^d$. The energy functional is the entropy.

$$E(\rho) = \int \rho \log \rho - \rho.$$

The mobility is $m(\rho) = 1$ and the mass creation is penalized by the entropy, that is

$$e(h) = \int h \log h - h.$$

Let ρ and $\tilde{\rho}$ be generated via the minimizing movement scheme. Then in Section 2.6 of Chapter 2 we show that

$$\partial_s \rho = \Delta \rho - \rho,$$

$$\partial_s \tilde{\rho} = \Delta \tilde{\rho} - \tilde{\rho},$$

and

$$\mathcal{A}(\rho_s, \tilde{\rho}_s) \leq e^{-s}\mathcal{A}(\rho_0, \tilde{\rho}_0). \tag{1.6.52}$$

Here,

$$\mathcal{A}(\mu, \nu) := \inf C(v_t, h_{t,}),$$

and

$$C(v_t, h_t) = \int_0^1 \frac{1}{2} \int |v_t|^2 \rho_t \, dx + \int e(h_t) \, dx \, dt.$$

The infimum is taken among all positive measured valued maps from $[0, 1]$ to $\mathcal{M}(\mathbb{R}^d)$, satisfying $\rho_0 dx = \mu$, $\rho_1 dx = \nu$ and

$$\frac{d}{dt} \int \zeta \rho_t dx = \int \nabla \zeta v_t \rho_t dx - \int \zeta h_t dx, \quad \forall t \in [0, 1] \text{ and } \forall \zeta \in C^\infty(\overline{\Omega}).$$

As in Section 5.2.4 of the introduction, we obtain (1.6.52) by using a vector field along the curve of measures to induced a variation with fixed end points and obtain a Gronwall estimate on the action (see Section 3 of the introduction to the differential structure for path space).

Chapter 2

A new family of transportation costs with applications to reaction-diffusion and parabolic equations with boundary conditions

This paper introduces a family of transportation costs between non-negative measures. This family is used to obtain parabolic and reaction-diffusion equations with drift, subject to Dirichlet boundary condition, as the gradient flow of the entropy functional $\int_\Omega \rho \log \rho + V\rho + 1\, dx$. In [43], Figalli and Gigli study a transportation cost that can be used to obtain parabolic equations with drift subject to Dirichlet boundary condition. However, the drift and the boundary condition are coupled in that work. The costs in this paper allow the drift and the boundary condition to be detached.

Keywords: transportation distances, gradient flows, reaction-diffusion equations, boundary conditions.

2.1 Introduction

The use of optimal transport for the study of evolutionary equations has proven to be a powerful method in recent years. More precisely, one of the most surprising achievements of [66, 82, 83] has been that many evolution equations of the form

$$\frac{d}{dt}\rho(t) = \mathrm{div}\Big(\nabla\rho(t) + \rho(t)\nabla V + \rho(t)(\nabla W * \rho(t))\Big),$$

can be seen as gradient flows of some entropy functional on the space of probability measures with respect to the Wasserstein distance:

$$W_2(\mu,\nu) = \inf\left\{\sqrt{\int |x-y|^2\, d\gamma(x,y)} \quad : \quad \pi_{1\#}\gamma = \mu, \pi_{2\#}\gamma = \nu\right\}.$$

In addition to the fact that this interpretation allows one to prove entropy estimates and functional inequalities (see [93, 94] for more details on this area, which is still very active and in constant evolution), this point of view provides a powerful variational method to prove the existence of solutions to the above equations: given a time step $\tau > 0$, and an initial measure ρ_0, construct an approximate solution by iteratively minimizing

$$\rho \to \frac{W_2(\rho,\rho_n)^2}{2\tau} + \int \left[\rho\log\rho + \rho V + \frac{1}{2}\rho\big(W*\rho\big)\right] dx = L[\rho|\rho_n],$$

where ρ_n is a minimum for $L[\rho|\rho_{n-1}]$.

This approach will always produce solutions to parabolic equations with Neumann boundary conditions. More recently, Figalli and Gigli [43] introduced a distance among positive measures in an open domain Ω. Such a distance allows one to use this approach to build solutions to the problem:

$$\begin{cases} \frac{d}{dt}\rho(t) = \operatorname{div}\Big(\nabla\rho(t) + \rho(t)\nabla V\Big) \quad \text{in } \Omega, \\ \rho = e^{-V} \quad \text{on} \quad \partial\Omega, \end{cases}$$

in bounded domains. Note, however, that the boundary condition for ρ is decided by the drift term appearing in the equation. Our goal here is to decouple the equation and the boundary condition. Also, we want to allow for the presence of a reaction term. Hence, inspired by [43], we introduce a new family of transportation costs in a bounded open domain Ω. This family allows us to build weak solutions to

$$\begin{cases} \partial_t\rho = \operatorname{div}\Big(\nabla\rho(t) + \rho(t)\nabla V\Big) - F_x'(\rho) \quad \text{in} \quad \Omega, \\ \rho = \rho_D \quad \text{on} \quad \partial\Omega. \end{cases} \tag{2.1.1}$$

Here, F is a function on $[0,\infty) \times \overline{\Omega}$. We will use the notation $F_x := F(\,\cdot\,, x)$. Also, we denote the first and second partial derivatives with respect to the first variable by F_x' and F_x''. Our method works for a wide class of reaction terms F_x'. Some examples include

$$F_x'(\rho) = W(x)\rho^{1+\beta} - Q(x),$$

$$F_x'(\rho) = W(x)\log\rho - Q(x),$$

and

$$F_x'(\rho) = W(x)(\rho - 1)|1 - \rho|^{\alpha - 1} - Q(x),$$

with α in $(0,1)$, $\beta \geq 0$, W Lipschitz and strictly positive, and Q Lipschitz and non-negative. (Note that when $V = W = 1$ and $Q = 0$, the last example is equivalent to the equation $\partial_t u = \Delta u - u^\alpha$ via the change of variable $u = \rho - 1$, for non negative initial data.)

Now, we list sufficient conditions on F :

(F1) F_x is strictly convex for every x in $\overline{\Omega}$.

(F2) For every x in $\overline{\Omega}$, F_x' is a homeomorphism from $(0,\infty)$ to $(\inf_{r>0} F_x'(r), \infty)$.

(F3) For every r in $(0,\infty)$ the map F_x' is a continuous function of x.

(F4) $\lim_{r\to\infty}[F_x'](r) = \infty$ uniformly in x.

(F5) There exist positive constants s, s_1, B_0, and C_0 such that,

$$F_x'(r) \leq C_0\, r,$$

for every (r,x) in $(0,s) \times \overline{\Omega}$ and

$$||\nabla_x[F_x'](e^{p-V(x)})||_{L^\infty(\Omega)} \leq B_0,$$

for every (p,x) in $(-\infty, -s_1) \times \overline{\Omega}$.

(F6) The map

$$(h,x) \to \int_{[F_x']^{-1}(0)}^{[F_x']^{-1}(h)} (\log r + V) F_x''(r)dr,$$

is Lipschitz on any compact subset of $\{(h,x) \in \mathbb{R} \times \overline{\Omega} \; : \; [F_x']^{-1}(h(x)) > 0\}$.

(F7) For every x in Ω, F'_x satisfies that either

$$\lim_{r\to 0} F'_x(r) = -\infty,$$

or

$$\lim_{r\downarrow 0} F'_x(r) = F'_x(0).$$

We will assume that the drift, the domain, and the boundary data satisfy:

(B1) V is Lipschitz.

(B2) Ω is Lipschitz, open, bounded, and satisfies the interior ball condition.

(B3) ρ_D is Lipschitz and uniformly positive.

These transportation costs, that we shall define later, were found through a set of heuristic arguments (see Section 2). These arguments explore costs that are related to a larger class of problems. Examples of these problems include:

$$\begin{cases} \partial_t \rho = \operatorname{div}\Big(\nabla\rho(t) + \rho(t)\nabla V\Big) - F'_x(\rho)m(\rho) \quad \text{in} \quad \Omega, \\ \rho = \rho_D \quad \text{on} \quad \partial\Omega, \end{cases} \tag{2.1.2}$$

and

$$\begin{cases} \partial_t \rho = \operatorname{div}\Big(\nabla\rho(t) + \rho(t)\nabla V\Big) - F'_x(\rho)m(\rho) \quad \text{in} \quad \Omega, \\ -\langle \nabla\rho - \nabla V\rho, \nu\rangle = g_R(\rho - \rho_R) \quad \text{on} \quad \partial\Omega. \end{cases} \tag{2.1.3}$$

Here, the functions g_R, and ρ_R are assumed to be uniformly positive. Also, $m : [0,\infty) \to [0,\infty)$ is concave.

The author found this heuristic by combining several previous works. First, the work of Felix Otto on the formal Riemannian structure in the space of probability measures [84, Section 3]. Second, the work of John Milnor [78, Part III] on the formal Riemannian structure in the space of paths of a Riemannian manifold. Third, the work of Francesco Rossi and Benedetto Piccoli [86] on the generalization of the Benamou Brenier formula [11] for positive measures. The last ingredient is a paper by Figalli, Gangbo, and Yolcu [42], in which they successfully follow the minimizing movement scheme for Lagrangian cost. The addition of nonlinear mobilities and the corresponding notion of generalized geodesics has been studied in a diffrent context by J.A Carrillo, S. Lisini, G. Savare, and D. Slepcev [21].

The heuristic arguments are developed in the second section of this paper. These are made rigorous only for the costs induced by Problem 1.1. These costs produce solutions to (2.1.1) and (2.1.7) via the minimizing movement scheme. This is the main result of the paper: Theorem 2.4.1.

Our family of costs depend on a positive number τ and two functions

$$e \ : \ \mathbb{R} \times \overline{\Omega} \to \mathbb{R} \cup \{\infty\},$$

and

$$\Psi \ : \ \overline{\Omega} \to \mathbb{R}.$$

We will use the notation $e_x := e(\,\cdot\,, x)$. We will denote the derivative of e with respect to its first entry by e'_x. Additionally, for each fixed x, $[e'_x]^{-1}$

denotes the inverse of such a derivative as a function of its own first entry. Analogous notation will be used for F. We will denote the interior of the set of points such that e is finite by $D(e)$ and the interior of the set of points z such that $e(z,x)$ is finite by $D(e_x)$. We require that the functions Ψ and e satisfy the following properties:

(C1) Ψ is Lipschitz.

(C2) For each x in $\overline{\Omega}$, $e_x := e(\,\cdot\,,x)$ is strictly convex and lower semicontinuous.

(C3) For each $L \in \mathbb{R}$, there exists $C(L)$ such that

$$e(z,x) \geq L|z| + C(L) \quad \forall (z,x) \in \mathbb{R} \times \overline{\Omega}.$$

(C4) The map e is Lipschitz in any compact subset of $D(e)$. (We regard $\overline{\Omega}$ as a topological space: Hence, the interior of any set of the form $\overline{A} \times \overline{\Omega}$, where A is an open subset of $\mathbb{R}$, is given by $A \times \overline{\Omega}$).

(C5) For each x in Ω, the sets $D(e_x)$ are of the form $(a(x),\infty)$, with $a(x)$ being either a constant or negative infinity.

(C6) For each x, the map e'_x is a homeomorphism between $D(e_x)$ and $\mathbb{R}$.

(C7) For each r in $\mathbb{R}$, the map $[e'_x]^{-1}(r)$ is a continuous function of x and

$$\lim_{r\to\infty} [e'_x]^{-1}(r) = \infty,$$

uniformly in x.

(C8) There exist positive constants s, s_1, B_0, and C_0 such that

$$[e'_x]^{-1}\big(\log r + V(x)\big) \leq C_0\, r,$$

for every (r,x) in $(0,s)\times\overline{\Omega}$ and

$$||\nabla_x [e'_x]^{-1}(p)||_{L^\infty(\Omega)} \leq B_0,$$

for every (p,x) in $(-\infty, s_1)\times\overline{\Omega}$.

(C9) The function e statisfies that

$$\int_\Omega e(0,x)\,dx = 0.$$

Item (C9) can be easily be relaxed by adding a constant to e; we have just assumed it for convenience. The notations $e(h(x),x)$, $e(h)$, $e\circ h$, and $e_x(h)$ will be used interchangeably. Similarly, we will freely interchange $e'(h(x),x)$, $e'(h)$, $e'_x(h)$, and $e'\circ h$.

We will use Ψ to obtain the desired boundary condition and e to control the reaction term. We define the cost $Wb_2^{e,\Psi,\tau}$ on the set of positive measures with finite mass $\mathcal{M}(\Omega)$, as a result of Problem 1.1, below.

Problem 1.1 (A variant of the transportation problem). *Given $\mu, \rho\, dx \in \mathcal{M}(\Omega)$, we consider the problem of minimizing*

$$C_\tau(\gamma,h) := \int_{\overline{\Omega}\times\overline{\Omega}\setminus\partial\Omega\times\partial\Omega}\left(\frac{1}{2}\frac{|x-y|^2}{\tau} + \Psi(y)1_{\Omega\times\partial\Omega} - \Psi(x)1_{\partial\Omega\times\Omega}\right)d\gamma \\ + \tau\int_\Omega e(h)\,dx, \quad (2.1.4)$$

in the space $ADM(\mu, \rho)$ of admissible pairs (γ, h). An admissible pair consists of a positive measure γ in $\overline{\Omega} \times \overline{\Omega}$ and a function h in $L^1(\Omega)$. We require the pair to satisfy

$$\pi_{2\#}\gamma_{\Omega}^{\Omega} = \rho\, dx + \tau h\, dx \quad \textit{and} \quad \pi_{1\#}\gamma_{\Omega}^{\overline{\Omega}} = \mu. \tag{2.1.5}$$

Here, the measure γ_A^B denotes the restriction of γ to $A \times B \subset \overline{\Omega} \times \overline{\Omega}$. Also, the functions π_1 and π_2 are the canonical projections of $\overline{\Omega} \times \overline{\Omega}$ into the first and second factor.

Hence, (2.1.4) provides a transportation cost between μ and ρ given by

$$Wb_2^{e,\Psi,\tau}(\mu, \rho) := \inf_{(\gamma,h) \in ADM} C_\tau(\gamma, h).$$

Additionally, we will denote by $Opt(\mu, \nu)$ the set of minimizers of Problem 1.1 with μ and ν given.

The main objective is the following: given an initial measure ρ_0, we build a family of curves $t \to \rho^\tau(t)$, indexed by $\tau > 0$. We will do this by iteratively minimizing

$$\rho \to \int_\Omega [\rho \log \rho - \rho + V(x)\rho + 1]\, dx + Wb_2^{e,\Psi,\tau}(\rho_n^\tau, \rho) = E^\tau[\rho|\rho_n^\tau], \tag{2.1.6}$$

where ρ_n^τ is a minimum of $E^\tau[\rho|\rho_{n-1}^\tau]$ in $\mathcal{M}(\Omega)$. We define the discrete solutions by

$$\rho^\tau(t) := \rho^\tau_{[t/\tau]}.$$

We then show that as $\tau \downarrow 0$, we can extract a subsequence converging to a weak solution to the problem:

$$\begin{cases} \partial_t \rho = \mathrm{div}\Big(\nabla\rho + \rho\nabla V\Big) - [e_x']^{-1}(\log\rho + V), \quad \text{in} \quad \Omega, \\ \rho = e^{\Psi - V}, \quad \text{in} \quad \partial\Omega, \\ \rho(0) = \rho_0. \end{cases} \tag{2.1.7}$$

In particular when we set

$$e(z,x) = \begin{cases} \int_{[F_x']^{-1}(0)}^{[F_x']^{-1}(z)} \big(\log r + V(x)\big) F_x''(r)\, dr, \quad \text{if} \quad z > \inf_{r>0} F_x'(r), \\ \liminf_{z \downarrow F_x'(0)} \int_{[F_x']^{-1}(0)}^{[F_x']^{-1}(z)} \big(\log r + V(x)\big) F_x''(r)\, dr, \quad \text{if} \quad z = \inf_{r>0} F_x'(r), \\ +\infty \quad \text{otherwise}, \end{cases} \tag{2.1.8}$$

and

$$\Psi = \log \rho_D + V \quad \text{on} \quad \partial\Omega, \tag{2.1.9}$$

we obtain a weak solution to (2.1.1).

Whenever the reaction term satisfies (F1)-(F7), the drift, the boundary, and the boundary data satisfy (B1)-(B3), and Ψ and e are as above, then properties (C1)-(C9) are satisfied as well.

We will require ρ_0 to be bounded and uniformly bounded away from zero. Using Proposition 2.5.2, we will show the existence of positive constants λ and Λ such that the weak solution satisfies

$$\frac{\lambda}{\sup e^{-V}} e^{-(C_0 t + V)} \le \rho(x,t) \le \frac{\Lambda}{\inf e^{-V}} e^{-V},$$

for almost every x.

The paper is organized as follows: Section 2 introduces the heuristics used to find the transportation costs. There, we explain the process used to relate the cost with the boundary conditions and reaction term. Section 3 is devoted to the study of Problem 1.1 and characterization of its solutions in terms of convex functions. Section 4 is devoted to the proof of the main result, Theorem 2.4.1, which states the convergence of the minimizing movement scheme to the weak solution. Section 5 is devoted to the study of properties of the minimizers of $E^\tau[\rho|\rho_0]$ that we use to prove the main Theorem. Finally, Appendix A is used to prove some technical properties of solutions to Problem 1.1 that are necessary in Section 3.

Acknowledgments: All the works used to find the new heuristics, with the exception of the work of Milnor, were introduced to me by my PhD advisor Alessio Figalli. I want to express my gratitude to him for his invaluable support and orientation. The work of Milnor was introduced to me by my undergraduate advisor Lazaro Recht. I also wish to express my gratitude to him for his invaluable support and orientation.

2.2 Heuristics

We define the cost $Wb_2^{e,m,\overline{e},\tau}$, as a result of Problem 2.1, below.

Problem 2.1 (A variant of the transportation problem). *Given $\mu, \nu \in \mathcal{M}(\Omega)$ we consider the problem of minimizing*

$$\tilde{C}_\tau(v_t, h_{t,}, \overline{h_t}) = \int_0^\tau \left[\frac{1}{2}\int_\Omega |v_t|^2 \rho_t \, dx + \int_\Omega e(h_t) m(\rho_t) \, dx + \int_{\partial\Omega} \overline{e}(\overline{h}_t) \, d\mathcal{H}^{d-1}\right] dt,$$

among all positive measured valued maps from $[0,\tau]$ to $\mathcal{M}(\Omega)$, satisfying $\rho_0 \, dx = \mu$ and $\rho_\tau \, dx = \nu$. Here, the measures ρ_t and the triplets $(v_t, h_{t,}, \overline{h_t})$ are indexed by t in $[0,\tau]$. We require them to satisfy the constraint

$$\frac{d}{dt}\int_\Omega \zeta \rho_t \, dx = \int_\Omega \langle \nabla\zeta, v_t\rangle \rho_t \, dx - \int_\Omega \zeta h_t m(\rho) \, dx - \int_{\partial\Omega} \zeta \overline{h}_t \, d\mathcal{H}^{d-1}, \quad \forall t \in [0,\tau] \text{ and } \forall \zeta \in C^\infty(\overline{\Omega}). \quad (2.2.10)$$

This provides a transportation cost between μ and ν given by

$$Wb_2^{e,m,\overline{e},\tau}(\mu,\nu) := \inf \tilde{C}_\tau(v_t, h_{t,}, \overline{h_t}).$$

Henceforth, a path is defined as a measured valued map from $[0,\tau]$ to $\mathcal{M}(\Omega)$. We apply the minimizing movement scheme to this cost: given an initial measure

ρ_0, we build a family of curves $t \to \rho^\tau(t)$, indexed by $\tau > 0$, iterating the minimization of the map

$$\rho \to \int_\Omega [\rho \log \rho - \rho + V(x)\rho + 1]\, dx + Wb_2^{e,m,\overline{e},\tau}(\rho_n^\tau, \rho) = \tilde{E}^\tau[\rho|\rho_n^\tau],$$

where ρ_n^τ is a minimum of $\tilde{E}^\tau[\rho|\rho_{n-1}^\tau]$, in $\mathcal{M}(\Omega)$. We define the discrete solutions by

$$\rho^\tau(t) := \rho_{[t/\tau]}^\tau.$$

Then, as $\tau \downarrow 0$, we extract a subsequence converging to a weak solution of the problem:

$$\begin{cases} \partial_t \rho = \operatorname{div}\Big(\nabla\rho + \rho\nabla V\Big) - [e_x']^{-1}(\log\rho + V)m(\rho), \quad in \quad \Omega, \\ -\langle \nabla\rho - \nabla V\rho, \nu\rangle = [\overline{e}_x']^{-1}(\log\rho + V) \quad in \quad \partial\Omega, \\ \rho(0) = \rho_0. \end{cases}$$

In particular, when we set

$$e(h,x) = \begin{cases} \int_{[F_x']^{-1}(0)}^{[F_x']^{-1}(h)} (\log r + V)F_x''(r)\, dr, \quad \text{if} \quad [F_x']^{-1}(h) \geq 0, \\ +\infty, \quad \text{otherwise}, \end{cases}$$

and

$$\overline{e}(\overline{h},x) = \begin{cases} g_R\Big(l(\overline{h})\log l(\overline{h}) + (V-1)l(\overline{h}) + 1\Big), \quad \text{if} \quad l(\overline{h}) \geq 0, \\ +\infty, \quad \text{otherwise}, \end{cases}$$

we obtain a weak solution to: problem (2.1.3). Here,

$$l(r) = \frac{r}{g_R} + \rho_R.$$

Also, we will show that when we set

$$e(h,x) = \begin{cases} \int_{[F_x']^{-1}(0)}^{[F_x']^{-1}(h)} (\log r + V)F_x''(r)\, dr, \quad \text{if} \quad [F_x']^{-1}(h) \geq 0, \\ +\infty, \quad \text{otherwise}, \end{cases}$$

and

$$\overline{e}(\overline{h}, x) = (\log \rho_D + V)\overline{h},$$

we obtain a weak solution to (2.1.2).

The heuristic is presented as follows. Section 2.1 characterizes optimal triplets in terms of potentials. Section 2.2 describes a characterization of minimal paths in terms of an equation for the potentials. Section 2.3 describes how the equation for minimal paths can be used to perform the minimizing movement scheme. Section 2.4 describes how to match the cost with the boundary conditions. Finally, section 2.5 describes how to match the cost with the reaction term.

2.2.1 Optimal triplets

In this section, we show a heuristic argument that characterizes minimizing triplets for Problem 2.1. For such triplets, there exist functions φ_t indexed in $[0, \tau]$, such that:

(a) $\nabla\varphi_t = v_t$.

(b) $\varphi_t = -\overline{e}'(\overline{h}_t)$ on $\partial\Omega$ and $\overline{h_t} = \langle \rho\nabla\varphi_t, \nu\rangle$.

(c) $\varphi_t = -e'(h_t) \quad \text{in} \quad \Omega$.

In order to see this, we fix $t \in [0, \tau]$ and minimize

$$\frac{1}{2}\int_\Omega |v_t|^2\rho_t \; dx + \int_\Omega e(h_t)m(\rho_t)\, dx + \int_{\partial\Omega} e(\overline{h_t})\, d\mathcal{H}^{d-1},$$

under the constraint (2.2.10).

First, we prove (a).

Let us assume that we have a minimizer for Problem 2.1. Let $(v_t, h_t, \overline{h_t})$ be the corresponding minimal triplet at the given time. We proceed as in the classical case, [5, Proposition 2.30]. Let W be a compactly supported vector field in the interior of Ω, with $\operatorname{div}(\rho_t W) = 0$. Then, $(v_t + sW, h_t, \overline{h_t})$ still satisfies the constraint, for every s. Hence, by minimality, we must have

$$\frac{d}{dt}\bigg|_{s=0} \frac{1}{2}\int_\Omega |v_t+sW|^2\rho_t\,dx+\int_\Omega e(h_t)m(\rho_t)\,dx+\int_{\partial\Omega}\overline{e}(\overline{h_t})\,d\mathcal{H}^{d-1} = \int_\Omega \langle v_t, W\rangle\rho_t\,dx = 0.$$

Since W was an arbitrary vector field satisfying $\operatorname{div}(W\rho_t) = 0$, by the Helmholtz-Hodge Theorem we obtain $\nabla\varphi_t = v_t$ for some $\varphi_t : \Omega \to \mathbb{R}$.

Second, we prove (b).

Let $\omega : \partial\Omega \to \mathbb{R}$ be a smooth function. Also, let α solve the elliptic problem

$$\begin{cases} \operatorname{div}(\rho_t\nabla\alpha) = 0 \quad in \quad \Omega, \\ \langle \rho_t\nabla\alpha, \nu\rangle = \omega \quad in \quad \partial\Omega. \end{cases} \tag{2.2.11}$$

Then, $(v_t + s\nabla\alpha, , h_t, \overline{h}_t + s\omega)$ satisfies the constraint for any s. By minimality, we must have

$$\frac{d}{ds}\bigg|_{s=0} \frac{1}{2}\int_\Omega |\nabla\varphi_t + s\nabla\alpha|^2\rho_t\,dx + \int_\Omega e(h_t)m(\rho_t)\,dx + \int_{\partial\Omega}\overline{e}(\overline{h}_t + s\omega)\,d\mathcal{H}^{d-1} = 0.$$

Hence,

$$\int_\Omega \langle\nabla\varphi_t, \nabla\alpha\rangle\rho_t\,dx + \int_{\partial\Omega}\omega\overline{e}'(\overline{h_t})\,d\mathcal{H}^{d-1} = 0.$$

Integrating by parts and using (2.2.11), we obtain

$$\int_{\partial\Omega}\omega(\overline{e}'(\overline{h_t}) + \varphi_t)d\mathcal{H}^{d-1} = 0.$$

Since ω was arbitrary, we conclude

$$\overline{e}'(\overline{h}_t) = -\varphi_t \quad \text{on} \quad \partial\Omega.$$

By (2.2.10), we must have

$$\begin{aligned}\int_\Omega \zeta\partial_t\rho_t\,dx &= \int_\Omega \langle\nabla\zeta,\nabla\varphi_t\rangle\rho_t\,dx - \int_\Omega \zeta h_t\rho\,dx - \int_{\partial\Omega}\zeta\overline{h_t}\rho_t\,d\mathcal{H}^{d-1}\\ &= -\int_\Omega \zeta\operatorname{div}(\nabla\varphi\rho_t) - \int_\Omega \zeta h_t\rho\,dx + \int_{\partial\Omega}\zeta\Big(\langle\nabla\varphi\rho_t,\nu\rangle - \overline{h_t}\Big)d\mathcal{H}^{d-1},\end{aligned}$$

for any $\zeta : \Omega \to \mathbb{R}$.

Thus, we conclude

$$\langle\nabla\varphi\rho_t,\nu\rangle = \overline{h}_t \quad \text{on} \quad \partial\Omega.$$

Third, we show (c).

Let β, $\eta : \Omega \to \mathbb{R}$ be smooth compactly supported functions satisfying

$$-\operatorname{div}(\nabla\beta\rho_t) = m(\rho_t)\eta. \tag{2.2.12}$$

Then, for any s, the triplet $(\nabla\varphi + s\nabla\beta, h + s\eta, \overline{h})$ is admissible. Consequently, we must have

$$\frac{d}{ds}\bigg|_{s=0}\frac{1}{2}\int_\Omega |\nabla\varphi_t + s\nabla\beta|^2\rho_t\,dx + \int_\Omega e(h_t+s\eta)m(\rho_t)\,dx + \int_{\partial\Omega}\overline{e}(\overline{h}_t)\,d\mathcal{H}^{d-1} = 0.$$

Hence,

$$\int_\Omega \langle\nabla\varphi_t,\nabla\beta\rangle\rho_t\,dx + \int_\Omega \eta\, e'(h_t)m(\rho_t)\,dx = 0.$$

Integrating by parts and using (2.2.12), we obtain

$$\int_\Omega [\varphi_t + e'(h_t)]\eta m(\rho_t)\,dx = 0.$$

Since η was arbitrary, we conclude

$$e'(h_t) = -\varphi_t \quad \text{in} \quad \Omega.$$

2.2.2 Optimal paths

In this section, we will show a heuristic argument that characterizes minimizers of Problem 2.1.

Let ρ_t, indexed in $[0, \tau]$, be a minimizer of Problem 2.1. Also, for each t in $[0, \tau]$, let φ_t be the potential generating the corresponding optimal triplet: $\big(\nabla\varphi_t, [e']^{-1}(-\varphi_t), [\overline{e}']^{-1}(-\varphi_t)\big)$. Then,

$$\partial_t \varphi + \frac{1}{2}|\nabla\varphi_t|^2 - \big[\varphi_t[e']^{-1}(-\varphi_t) + e([e']^{-1}(-\varphi_t))\big]m'(\rho_t) = 0, \tag{2.2.13}$$

and

$$\frac{d}{dt}\int_\Omega \zeta\rho_t\, dx = \int_\Omega \langle\nabla\zeta, \nabla\varphi_t\rangle\rho_t\, dx - \int_\Omega \zeta[e']^{-1}(-\varphi_t)m(\rho_t)\, dx \\ - \int_{\partial\Omega} \zeta[\overline{e}']^{-1}(-\varphi_t)\, d\mathcal{H}^{d-1},$$

for every ζ in $C_c^\infty(\overline{\Omega})$.

In order to see this, we proceed by perturbing such minimizers. For each $t \in [0, \tau]$, we consider optimal triplets $\big(\nabla\omega_t, [e']^{-1}(-\omega_t), [\overline{e}']^{-1}(-\omega_t)\big)$. We require ω_t to be identically 0 in the complement of a compact subset of $(0, \tau)$.

Then, for each s, we let $t \to \rho_{t,s}$ and

$$t \to \big(\nabla\varphi_{t,s}, [e']^{-1}(-\varphi_{t,s}), [\overline{e}']^{-1}(-\varphi_{t,s})\big)$$

satisfy constraint (2.2.10). Additionally, for each t, we require the map $s \to \rho_{t,s}$ to satisfy

$$\frac{d}{ds}\bigg|_{s=0} \int_\Omega \zeta \rho_{t,s}\, dx = \int_\Omega \langle \nabla \zeta, \nabla \omega_t \rangle \rho_{t,s}\, dx - \int_\Omega \zeta [e']^{-1}(-\omega_t) m(\rho_{t,s})\, dx \\ - \int_{\partial\Omega} \zeta [\overline{e}']^{-1}(-\omega_t)\, d\mathcal{H}^{d-1},$$

and

$$\rho_{t,0} = \rho_t, \quad \varphi_{t,0} = \varphi_t.$$

Since $t \to \rho_t$ is a minimizer, we must have

$$\frac{d}{ds}\bigg|_{s=0} \int_0^\tau \frac{1}{2}\bigg[\int_\Omega |\nabla \varphi_{t,s}|^2 \rho_{t,s}\, dx + \int_\Omega e([e']^{-1}(-\varphi_{t,s})) m(\rho_{t,s})\, dx \\ + \int_{\partial\Omega} \overline{e}([\overline{e}']^{-1}(-\varphi_{t,s}))\, d\mathcal{H}^{d-1}\bigg] dt = 0.$$

Consequently,

$$\int_0^\tau \bigg[\int_\Omega \langle \nabla \varphi_{t,s}, \nabla \partial_s \varphi_{t,s} \rangle \rho_{t,s}\, dx + \frac{1}{2} \int_\Omega |\nabla \varphi_{t,s}|^2 \partial_s \rho_{t,s}\, dx \\ - \int_\Omega \varphi_{t,s} \partial_s [e']^{-1}(-\varphi_{t,s}) m(\rho_{t,s})\, dx \\ + \int_\Omega e([e']^{-1}(-\varphi_{t,s})) m'(\rho_{t,s}) \partial_s \rho_{t,s}\, dx - \int_{\partial\Omega} \varphi_{t,s} \partial_s [\overline{e}']^{-1}(-\varphi_{t,s})\, d\mathcal{H}^{d-1}\bigg] dt = 0.$$

Then,

$$\begin{aligned}\int_0^\tau \Bigg[\frac{d}{ds}\Bigg(\int_\Omega |\nabla\varphi_{t,s}|^2\rho_{t,s}\,dx - \int_\Omega \varphi_{t,s}[e']^{-1}(-\varphi_{t,s})m(\rho_{t,s})\,dx \\ - \int_{\partial\Omega}\varphi_{t,s}[\overline{e}']^{-1}(-\varphi_{t,s})\,d\mathcal{H}^{d-1}\Bigg) \\ - \Bigg(\int_\Omega \langle\nabla\partial_s\varphi_{t,s},\nabla\varphi_{t,s}\rangle\rho\,dx - \int_\Omega \partial_s\varphi_{t,s}[e']^{-1}(-\varphi_{t,s})m(\rho_{t,s})\,dx \\ - \int_{\partial\Omega}\partial_s\varphi_{t,s}[\overline{e}']^{-1}(-\varphi_{t,s})\,d\mathcal{H}^{d-1}\Bigg) - \frac{1}{2}\int_\Omega |\nabla\varphi_{t,s}|^2\partial_s\rho_{t,s}\,dx \\ + \int_\Omega \varphi_{t,s}[e']^{-1}(-\varphi_{t,s})m'(\rho_{t,s})\partial_s\rho_{t,s}\,dx \\ + \int_\Omega e([e']^{-1}(-\varphi_{t,s}))m'(\rho_{t,s})\partial_s\rho_{t,s}\,dx\Bigg]dt = 0.\end{aligned}$$

Recall that $h_{t,s} = [e']^{-1}(-\varphi_{t,s})$ and $\overline{h}_{t,s} = [\overline{e}']^{-1}(-\varphi_{t,s})$. By (2.2.10), we get

$$\begin{aligned}\int_0^\tau \Bigg[\frac{d}{ds}\int_\Omega \varphi_{t,s}\partial_t\rho_{t,s}\,dx - \int_\Omega \partial_s\varphi_{t,s}\partial_t\rho_{t,s}\,dx \\ - \int_\Omega \Bigg(\frac{1}{2}|\nabla\varphi_{t,s}|^2 - \big[\varphi_{t,s}[e']^{-1}(-\varphi_{t,s}) \\ + e([e']^{-1}(-\varphi_{t,s}))\big]m'(\rho_{t,s})\Bigg)\partial_s\rho_{t,s}\,dx\Bigg]dt = 0. \quad (2.2.14)\end{aligned}$$

By construction $\partial_s\varphi_{\tau,s} = \partial_s\rho_{\tau,s} = \partial_s\varphi_{0,s} = \partial_s\rho_{0,s} = 0$. Hence, if we integrate by parts in t, we obtain

$$\begin{aligned}-\int_0^\tau \Bigg[\int_\Omega \Bigg(\partial_t\varphi + \frac{1}{2}|\nabla\varphi_t|^2 - \big[\varphi_t[e']^{-1}(-\varphi_t) \\ + e([e']^{-1}(-\varphi_t))\big]m'(\rho_{t,s})\Bigg)\partial_s\rho_{t,s}\,dx\Bigg]dt = 0. \quad (2.2.15)\end{aligned}$$

This gives the desired result.

2.2.3 The minimizing movement scheme.

Given $\rho_0 \in \mathcal{M}(\Omega)$ and $\tau > 0$, we provide heuristic arguments to characterize the minimizers of

$$\{\rho_t\}_{t\in[0,\tau]} \to \int_0^\tau \left(\frac{1}{2}\int_\Omega |v_t|^2\rho_t\,dx + \int_\Omega e(h_t)m(\rho_t)\,dx + \int_{\partial\Omega} \overline{e}(\overline{h}_t)\,d\mathcal{H}^{d-1}\right)dt \\ + \int_\Omega \big[\rho_\tau \log\rho_\tau + \big(V-1\big)\rho_\tau + 1\big]\,dx. \tag{2.2.16}$$

Here, the triplets $(v_t, h_t, \overline{h}_t)$ satisfy (2.2.10). Also, ρ_0 is fixed and $\rho_\tau = \rho$. In Section 2.1, we saw that for minimizing triplets we have for each $t \in [0,\tau]$ a function φ_t, such that

$$(v_t, h_t, \overline{h}_t) = \big(\nabla\varphi_t, [e']^{-1}(-\varphi_t), [\overline{e}']^{-1}(-\varphi_t)\big).$$

In Section 2.2, we found that optimal paths satisfy

$$\partial_t\varphi_t + \frac{1}{2}|\nabla\varphi_t|^2 - \big[\varphi_t[e']^{-1}(-\varphi_t) + e([e']^{-1}(-\varphi_t))\big]m'(\rho_t) = 0.$$

In this section, we will show that minimizers of (2.2.16) must satisfy

$$\varphi_\tau = -\log\rho_\tau - V.$$

In order to see this, we suppose that we have a minimizer ρ_τ and a path $t \to \rho_t$, with corresponding triplets $t \to \big(\nabla\varphi_t, [e']^{-1}(-\varphi_t), [\overline{e}']^{-1}(-\varphi_t)\big)$.

We proceed by perturbing the path $t \to \rho_t$. For each t we choose a function ω_t. We require these functions to be identically 0 in the complement of a compact subset of $(0,\tau]$. This generates for each s a path $t \to \rho_{t,s}$, as in Section 2.2, with the difference that now the end point $\rho_{\tau,s}$ is free.

For a minimizer, we must have

$$\begin{aligned}\frac{d}{ds}\bigg|_{s=0}\bigg(\int_0^\tau \frac{1}{2}\int_\Omega |\nabla\varphi_{t,s}|^2\rho_{t,s}\,dx \\ &+\int_\Omega e([e']^{-1}(-\varphi_{t,s}))m(\rho_{t,s})\,dx \\ &+\int_{\partial\Omega} \overline{e}([\overline{e}']^{-1}(-\varphi_{t,s}))\,d\mathcal{H}^{d-1}dt+\int_\Omega \big[\rho_{\tau,s}\log\rho_{\tau,s} \\ &+\big(V-1\big)\rho_{\tau,s}+1\big]\,dx\bigg)=0.\end{aligned}$$

By (2.2.14) we have

$$\begin{aligned}\int_0^\tau \bigg[\frac{d}{ds}\int_\Omega \varphi_{t,s}\partial_t\rho_{t,s}\,dx-\int_\Omega \partial_s\varphi_{t,s}\partial_t\rho_{t,s}\,dx \\ &-\int_\Omega\bigg(\frac{1}{2}|\nabla\varphi_{t,s}|^2-\big[\varphi_{t,s}[e']^{-1}(-\varphi_{t,s}) \\ &+e([e']^{-1}(-\varphi_{t,s}))\big]m'(\rho_{t,s})\bigg)\partial_s\rho_{t,s}\,dx\bigg]dt \\ &+\int_\Omega\big[\log\rho_\tau+V\big]\partial_s\rho_{\tau,s}\,dx=0.\end{aligned}$$

Then, if we use (2.2.13) we obtain

$$\begin{aligned}\int_0^\tau\bigg[\int_\Omega \varphi_{t,s}\partial_t\partial_s\rho_{t,s}\,dx \\ &+\int_\Omega \partial_t\varphi_{t,s}\partial_s\rho_{t,s}\,dx+\bigg]dt+\int_\Omega(\log\rho_\tau+V)\partial_s\rho_{\tau,s}\,dx=0.\end{aligned}$$

Recall that by construction $\partial_s\varphi_{0,s}=\partial_s\rho_{0,s}=0$. Integrating by parts we get

$$\int_\Omega\bigg(\varphi_\tau+\log\rho_\tau+V\bigg)\partial_s\rho_{\tau,s}\,dx=0.$$

Thus, we obtain the desired result.

2.2.4 The Boundary Conditions

In Section 2.3, we showed that for minimizers of (2.2.16), we have that $\varphi_\tau = -\log\rho_\tau - V$. In Section 2.1, we showed that for optimal triplets,

$$-\varphi = \overline{e}'(\overline{h}) \quad \text{on} \quad \partial\Omega.$$

Hence, if we set $\overline{e}(\overline{h}) = \Psi\,\overline{h}$, we obtain the boundary condition

$$\rho_\tau = e^{\Psi - V} \quad \text{on} \quad \partial\Omega.$$

This concludes the analysis for the boundary condition for (2.1.2).
In order to derive the boundary condition for (2.1.3), we proceed as follows: In Section 2.1, we showed that for minimizers of Problem 2.1, we must have

$$\overline{h}_t = \langle \rho_t \nabla\varphi_t, \nu\rangle.$$

Hence, we expect the limit of the minimizing movement scheme to satisfy the relation

$$-\langle\nabla\rho,\nu\rangle - \langle\rho\nabla V,\nu\rangle = [\overline{e}']^{-1}(\log\rho + V).$$

Our goal is to obtain the boundary condition

$$-\langle\nabla\rho,\nu\rangle - \langle\rho\nabla V,\nu\rangle = g_R\big(\rho - \rho_R\big).$$

For this purpose, we would need

$$[\overline{e}']^{-1}(\log\rho + V) = g_R\big(\rho - \rho_R\big).$$

Thus,

$$V + \log\rho = [\overline{e}']\bigg(g_R\big(\rho - \rho_R\big)\bigg).$$

Hence, if we set

$$l(r) = \frac{r}{g_R} + \rho_R,$$

we obtain

$$[\overline{e}'](l^{-1}(\rho)) = \log \rho + V.$$

Then, it follows that

$$[\overline{e}'](l^{-1}(\rho))[l^{-1}(\rho)]' = g_R(\log \rho + V).$$

Integrating, we obtain

$$\overline{e}(l^{-1}(\rho)) = \int_0^\rho g_R\big(\log r + V\big)dr + C.$$

Here, C is a constant that will be chosen later. This implies

$$\overline{e}(\rho) = g_R \int_{l(0)}^{l(\rho)} \big(\log r + V\big)dr + C.$$

Thus, it suffices to set

$$\overline{e}(\rho) = g_R\bigg(l(\rho)\log l(\rho) + \bigg(V - 1\bigg)l(\rho) + 1\bigg).$$

2.2.5 The reaction term

In Section 2.1, we showed that optimal triplets satisfy

$$e'(h_\tau) = -\varphi_\tau.$$

In Section 2.3, we showed that minimizers of (2.2.16) satisfy

$$\varphi_\tau = -\log \rho_\tau - V.$$

Thus, in order to obtain

$$h = F'(\rho),$$

we set

$$e'(F'(\rho)) = \log \rho + V.$$

This implies

$$e'(F'(\rho))F''(\rho) = \big(\log \rho + V\big)F''(\rho).$$

Integrating we obtain

$$e(F'(\rho)) = \int_0^\rho \big(\log r + V\big)F''(r)\,dr + C,$$

for some constant C. Thus, it suffices to set

$$e(\rho) = \int_{[F']^{-1}(0)}^{[F']^{-1}(\rho)} \big(\log r + V\big)F''(r)\,dr.$$

2.2.6 Stability estimates

Here we provide an example in which the minimizing movement scheme yields a contractible flow. By this we mean that the cost between any two solutions of the flow decreases exponentially in time. Here, $\Omega = \mathbb{R}^d$. The energy functional is the entropy.

$$E(\rho) = \int \rho \log \rho - \rho.$$

The mobility is $m(\rho) = 1$ and the mass creation is penalized by the entropy, that is

$$e(h) = \int h \log h - h.$$

Let ρ and $\tilde{\rho}$ be generated via the minimizing movement scheme. Then, one has

$$\partial_s \rho = \Delta \rho - \rho,$$

$$\partial_s \tilde{\rho} = \Delta \tilde{\rho} - \tilde{\rho},$$

and

$$\mathcal{A}(\rho_s, \tilde{\rho_s}) \leq e^{-s} \mathcal{A}(\rho_0, \tilde{\rho_0}).$$

Here,

$$\mathcal{A}(\mu, \nu) := \inf C(v_t, h_{t,}),$$

and

$$C(v_t, h_t) = \int_0^1 \frac{1}{2} \int |v_t|^2 \rho_t \, dx + \int e(h_t) \, dx \, dt.$$

The infimum is taken among all positive measured valued maps from $[0, 1]$ to $\mathcal{M}(\mathbb{R}^d)$, satisfying $\rho_0 dx = \mu$, $\rho_1 dx = \nu$ and

$$\frac{d}{dt} \int \zeta \rho_t dx = \int \nabla \zeta v_t \rho_t dx - \int \zeta h_t dx, \quad \forall t \in [0, 1] \text{ and } \forall \zeta \in C^\infty(\overline{\Omega}).$$

As in section 5.2.4 of the introduction, we obtain (1.6.52) by using a vector field along the curve of measures to induced a variation with fixed end points and obtain a Gronwall estimate on the action (see section 3 of the introduction to the differential structure for path space).

Proof.

Fix $s \geq 0$. Let $\rho_{s,t}$, with $t \in [0,1]$ be a path of measures minimizing $C(\rho_s, \tilde{\rho_s})$. Also, let (v_t, h_t) be the corresponding optimal triplet. By section 2 we know that there exist functions φ_t indexed in $[0,1]$ satisfying $(v_t, h_t) = (\nabla\varphi_t, e^{-\varphi_t})$. By section 3 we know

$$\partial_t \varphi + \frac{1}{2} \mid \nabla\varphi_t \mid^2 = 0,$$

$$\frac{d}{dt}\int \zeta\rho_{s,t} dx = \int \nabla\zeta\nabla\varphi\rho_{s,s} dx - \zeta e^{-\varphi} dx. \quad \forall\zeta \in C^\infty(\overline{\Omega}).$$

for all $t \in (0,1)$. We wish to perturb this path in the direction of the flow. To do this we assume that for each t

$$\frac{d}{ds}\int \zeta\rho_{s,t} dx = -\int \nabla\zeta\nabla \log\rho_{s,t}\rho_{s,t} dx - \int \zeta\rho_{s,t} dx \forall\zeta \in \quad C^\infty(\overline{\Omega}).$$

We wish to compute

$$\frac{d}{ds}\int_0^1 \left[\int \frac{1}{2} \mid \nabla\varphi \mid^2 \rho dx + \int \big(-\varphi e^{-\varphi} - e^{-\varphi} + 1\big) dx\right] dt.$$

By (2.2.14) the above expression is equal to

$$
\begin{aligned}
&\int_0^1 \left[\int \left(\frac{d}{ds}\varphi \partial_t \rho - \partial_s \varphi \partial_t \rho - \frac{1}{2} \mid \nabla \varphi \mid^2 \partial_s \rho \right) dx \right] dt \\
&\quad = \int_0^1 \left[\int \big(\varphi \partial_t \partial_s \rho + \partial_t \varphi \partial_s \rho \big) dx \right] dt \\
&\quad = \int_0^1 \left[\frac{d}{dt} \int \varphi \partial_s \rho dx \right] dt \\
&\quad = \int_0^1 \left[\frac{d}{dt} \Big(- \int \nabla \varphi \nabla \rho dx - \int \varphi \rho dx \Big) \right] dt \\
&\quad = \int_0^1 \left[\frac{d}{dt} \int \Big(\Delta \varphi - \varphi \Big) \rho dx \right] dt \\
&\quad = \int_0^1 \left[\int \left(\nabla \varphi \nabla \Delta \varphi - \frac{1}{2} \Delta \mid \nabla \varphi \mid^2 + \frac{\mid \nabla \varphi \mid^2}{2} - \mid \nabla \varphi \mid^2 \right) \rho dx \right. \\
&\quad \left. - \int \big(\Delta \varphi - \varphi \big) e^{-\varphi} dx \right] dt \\
&\quad = \int_0^1 \left[- \int \mathrm{Tr}(D^2 \varphi^T D^2 \varphi) \rho dx - \int \frac{1}{2} \mid \nabla \varphi \mid^2 \rho dx + \int \varphi e^{-\varphi} dx \right. \\
&\quad \left. - \int \mid \nabla \varphi \mid^2 e^{-\varphi} dx \right] dt \\
&\quad \leq - \int_0^1 \left[\int \frac{1}{2} \mid \nabla \varphi \mid^2 \rho dx + \int \varphi e^{-\varphi} dx + \int e^{-\varphi} dx \right] dt.
\end{aligned}
$$

Consequently,

$$
\begin{aligned}
&\frac{d}{ds} \int_0^1 \left[\int \frac{1}{2} \mid \nabla \varphi \mid^2 \rho dx + \int \big(- \varphi e^{-\varphi} - e^{-\varphi} \big) dx dt \leq \right. \\
&\qquad\qquad - \int_0^1 \left[\int \frac{1}{2} \mid \nabla \varphi \mid^2 \rho dx + \int \big(- \varphi e^{-\varphi} - e^{-\varphi} \big) dx \right] dt.
\end{aligned}
$$

Thus, the desired result follows by Gronwall inequality. This concludes the heuristic arguments. In the following sections we make these arguments rigorous for the case described in the introduction.

2.3 Properties of $Wb_2^{e,\Psi,\tau}$

In this section, we study minimizers of Problem 1.1. We begin by showing their existence.

Lemma 2.3.1. ***(Existence of Optimal pairs)*** *Let μ and ν be absolutely continuous measures in $\mathcal{M}(\Omega)$. Then there exists a minimizing pair for Problem 1.1.*

Proof. We claim the following:

Claim 1: There exists a minimizing sequence of admissible pairs

$$\{(\gamma_n, h_n)\}_{n=1}^{\infty}$$

for which the mass of $\{\gamma_n\}_{n=1}^{\infty}$ and $\{|h_n|\,dx\}_{n=1}^{\infty}$ is equibounded and the plans in the sequence have no mass concentrated on $\partial\Omega \times \partial\Omega$.

We assume this claim and postpone its proof until the end of the argument. By (2.1.5) the claim gives us a uniform bound in the total variation of $\{(\gamma_n, h_n\,dx)\}_{n=1}^{\infty}$. Then, by compactness of $\overline{\Omega}$ and $\overline{\Omega} \times \overline{\Omega}$, for a subsequence $\{(\gamma_n, h_n)\}_{n=1}^{\infty}$, not relabeled, we have weak convergence to regular Borel measures, with finite total variation, γ and $\mathbf{h}$. This converge is in duality with continuous bounded functions in $\overline{\Omega} \times \overline{\Omega}$ and $\overline{\Omega}$, respectively.

Assumption (C3) and the Dunford-Pettis Theorem allows us to conclude that $\mathbf{h} = h\,dx$, for some h in $L^1(\overline{\Omega})$ and that $\{h_n\}_{n=1}^{\infty}$ converges to h in duality with functions in $L^{\infty}(\overline{\Omega})$. Since $\pi_{2\#}(\gamma_n)_{\Omega}^{\Omega} = \rho\,dx + h_n\tau$, we have that for any $\zeta \in C_c(\overline{\Omega})$,

$$\int_{\overline{\Omega}\times\overline{\Omega}} \zeta \circ \pi_2 \, d\gamma_{\overline{\Omega}}^{\Omega} = \int_{\overline{\Omega}\times\overline{\Omega}} \zeta \circ \pi_2 \, d\gamma = \lim_{n\to\infty} \int_{\overline{\Omega}\times\overline{\Omega}} \zeta \circ \pi_2 \, d\gamma_n$$
$$= \lim_{n\to\infty} \int_{\overline{\Omega}\times\overline{\Omega}} \zeta \circ \pi_2 \, d(\gamma_n)_{\overline{\Omega}}^{\Omega}$$
$$= \lim_{n\to\infty} \int_{\overline{\Omega}} \zeta\rho \, dx + \tau \int_{\overline{\Omega}} \zeta h_n \, dx = \int_{\overline{\Omega}} \zeta\rho \, dx + \tau \int_{\overline{\Omega}} \zeta h \, dx.$$

Hence, $\pi_{2\#}\gamma_{\overline{\Omega}}^{\Omega} = \rho \, dx + h\tau$. It can also be shown that $\pi_{1\#}\gamma_{\Omega}^{\overline{\Omega}} = \mu$ in an analogous way. This implies that (γ, h) is in $ADM(\mu, \nu)$.

Since the sequence $\{h_n\}_{n=1}^{\infty}$ converges weakly in $L^1(\Omega)$ to h, using assumptions (C2)-(C6) and [64, Theorem 1], we get

$$\liminf_{n\to\infty} \int_{\Omega} e(h_n(x), x)dx \geq \int e(h(x), x)dx.$$

We also claim the following:

Claim 2: there exists a further subsequence $\{\gamma_n\}_{n=1}^{\infty}$, not relabeled, with the property that $\{(\gamma_n)_{\Omega}^{\Omega}\}_{n=1}^{\infty}$, $\{(\gamma_n)_{\partial\Omega}^{\Omega}\}_{n=1}^{\infty}$, and $\{(\gamma_n)_{\Omega}^{\partial\Omega}\}_{n=1}^{\infty}$ converge weakly to $\gamma_{\Omega}^{\Omega}, \gamma_{\partial\Omega}^{\Omega}$, and $\gamma_{\Omega}^{\partial\Omega}$ in duality with continuous and bounded functions in $C(\overline{\Omega} \times \overline{\Omega})$, $C(\partial\Omega \times \overline{\Omega})$, and $C(\overline{\Omega} \times \partial\Omega)$, respectively. We will also postpone the proof of this claim until the end of the argument.

Since Ψ is bounded and continuous, this claim implies that

$$\lim_{n\to\infty} \Bigg[\int_{\overline{\Omega}\times\overline{\Omega}} \frac{|x-y|^2}{2\tau} \, d(\gamma_n)_{\Omega}^{\Omega} + \int_{\partial\Omega\times\overline{\Omega}} \left(\frac{|x-y|^2}{2\tau} - \Psi(x) \right) d(\gamma_n)_{\partial\Omega}^{\Omega}$$
$$+ \int_{\overline{\Omega}\times\partial\Omega} \left(\frac{|x-y|^2}{2\tau} + \Psi(y) \right) d(\gamma_n)_{\Omega}^{\partial\Omega} \Bigg] = \int_{\overline{\Omega}\times\overline{\Omega}} \frac{|x-y|^2}{2\tau} \, d\gamma_{\Omega}^{\Omega}$$
$$+ \int_{\partial\Omega\times\overline{\Omega}} \left(\frac{|x-y|^2}{2\tau} - \Psi(x) \right) d\gamma_{\partial\Omega}^{\Omega} + \int_{\overline{\Omega}\times\partial\Omega} \left(\frac{|x-y|^2}{2\tau} + \Psi(y) \right) d\gamma_{\Omega}^{\partial\Omega}.$$

Hence, this shows the existence of minimizers, provided we prove the two claims. In order to prove the first one, we note that due to (2.1.4) and (2.1.5) we can assume, without loss of generality, that the plans in the minimizing sequence have no mass concentrated on $\partial\Omega\times\partial\Omega$. Also, due to (C3) and (2.1.5),

$$\begin{aligned} C_\tau(\gamma,h) &\geq -||\Psi||_\infty\Big(|\gamma^{\Omega}_{\partial\Omega}|+|\gamma^{\partial\Omega}_{\Omega}|\Big)+K\int_\Omega |h|\,dx + C(K)|\Omega| \\ &\geq -||\Psi||_\infty\Big(|\gamma^{\partial\Omega}_{\Omega}|+|\gamma^{\Omega}_{\partial\Omega}|\Big)+K|h|(\Omega)+C(K)|\Omega| \\ &\geq -||\Psi||_\infty\Big(\mu(\Omega)+\nu(\Omega)+\tau|h|(\Omega)\Big)+K|h|(\Omega)+C(K)|\Omega|, \end{aligned}$$

for any K. Taking K large enough, we obtain a uniform bound on $|h|(\Omega)$ and consequently on $|\gamma|$, for any minimizing sequence. This proves the first claim.

As previously explained, this claim gives us a subsequence, not relabeled $\{(\gamma_n,h_n)\}_{n=1}^\infty$, that converges weakly to (γ,h). To prove the second claim, we note that the measures in the sequence $\{(\gamma_n)^\Omega_\Omega,(\gamma_n)^\Omega_{\partial\Omega},(\gamma_n)^{\partial\Omega}_\Omega\}_{n=1}^\infty$ have uniformly bounded mass. Then, by compactness of $\overline{\Omega}\times\overline{\Omega}$, $\partial\Omega\times\overline{\Omega}$, and $\overline{\Omega}\times\partial\Omega$ we can find a further subsequence $\{(\gamma_n)^\Omega_\Omega,(\gamma_n)^\Omega_{\partial\Omega},(\gamma_n)^{\partial\Omega}_\Omega\}_{n=1}^\infty$, not relabeled, weakly converging to the measures σ_0,σ_1, and σ_2. This convergence is in duality with continuous and bounded functions in $C(\overline{\Omega}\times\overline{\Omega})$, $C(\partial\Omega\times\overline{\Omega})$, and $C(\overline{\Omega}\times\partial\Omega)$, respectively. Using the definition of weak convergence, it is easy to verify that we must have

$$\gamma=\sigma_0+\sigma_1+\sigma_2. \tag{2.3.17}$$

We will prove the second claim by showing that $\sigma_0 = \gamma_\Omega^\Omega$, $\sigma_1 = \gamma_{\partial\Omega}^\Omega$, and $\sigma_2 = \gamma_\Omega^{\partial\Omega}$. By (2.3.17), this is a consequence of the measures $\pi_{2\#}\sigma_0$, $\pi_{1\#}\sigma_0$, $\pi_{2\#}\sigma_1$, and $\pi_{1\#}\sigma_2$ having no mass concentrated in $\partial\Omega$. In order to see that these measures have this property, we let $A \subset \partial\Omega$ be a compact set and we take a sequence $\{\eta_k\}_{k=1}^\infty$ of uniformly bounded functions in $C(\overline{\Omega})$ that decreases monotonically to 1_A. Additionally, we require that the sets $supp(\eta_k)$ decrease monotonically to A. Since $\overline{\Omega}$ is bounded, by the dominated convergence Theorem,

$$\int_A d\pi_{2\#}\sigma_0 = \int_{\overline{\Omega}} 1_A \circ \pi_2 \, d\sigma_0 = \lim_{k\to\infty} \int_{\overline{\Omega}} \eta_k \circ \pi_2 \, d\sigma_0.$$

Also, by construction we have

$$\begin{aligned}\int_{\overline{\Omega}} \eta_k \circ \pi_2 \, d\sigma_0 &= \lim_{n\to\infty} \int_{\overline{\Omega}} \eta_k \circ \pi_2 \, d(\gamma_n)_\Omega^\Omega \le \lim_{n\to\infty} \int_{\overline{\Omega}\times\overline{\Omega}} \eta_k \circ \pi_2 \, d(\gamma_n)_{\overline{\Omega}}^\Omega \\ &= \lim_{n\to\infty} \int_{\overline{\Omega}} \eta_k \rho \, dx + \tau \int_{\overline{\Omega}} \eta_k h_n \, dx = \int_{\overline{\Omega}} \eta_k \rho \, dx + \tau \int_{\overline{\Omega}} \eta_k h \, dx \\ &= \int_{supp(\eta_k)} \eta_k (\rho + \tau h) \, dx \le \sup(\eta_k) \int_{supp(\eta_k)} |\rho + \tau h| \, dx.\end{aligned}$$

Since $\{\eta_k\}_{k=0}^\infty$ is uniformly bounded and $supp(\eta_k)$ converges monotonically to the set $A \subset \partial\Omega$ with zero $\mathcal{L}^d$ measure, we have

$$\int_A d\pi_{2\#}\sigma_0 \le \lim_{k\to\infty} \sup(\eta_k) \int_{supp(\eta_k)} |\rho + \tau h| \, dx = 0.$$

Thus, we conclude that $\pi_{2\#}\sigma_0(A) = 0$, for any measurable subset A of $\partial\Omega$; the proof for the measures $\pi_{1\#}\sigma_0, \pi_{2\#}\sigma_1$, and $\pi_{1\#}\sigma_2$ is analogous. This establishes the second claim. Consequently, the Lemma is proven. $\square$

We will use the following definitions:

Given an admissible pair (γ, h), we define

$$d_{\Psi,\tau}(x) = \begin{cases} \inf_{y\in\partial\Omega} \frac{|x-y|^2}{2\tau} + \Psi(y) \quad \text{if} \quad x \in \Omega, \\ 0 \quad \text{otherwise,} \end{cases}$$

and

$$d_{-\Psi,\tau}(y) = \begin{cases} \inf_{x\in\partial\Omega} \frac{|x-y|^2}{2\tau} - \Psi(x) \quad \text{if} \quad y \in \Omega, \\ 0 \quad \text{otherwise.} \end{cases}$$

For any x and y in Ω we denote by $\mathcal{P}_{\Psi,\tau}(x)$ and $\mathcal{P}_{-\Psi,\tau}(y)$ the sets where the infima are respectively attained. Henceforth, $P_{\Psi,\tau}$ and $P_{-\Psi,\tau}$ will be measurable maps from $\overline{\Omega}$ to $\partial\Omega$ such that

$$d_{\Psi,\tau}(x) = \frac{|x - P_{\Psi,\tau}(x)|^2}{2\tau} + 1_\Omega(x)\Psi\big(P_{\Psi,\tau}(x)\big),$$

and

$$d_{-\Psi,\tau}(y) = \frac{|y - P_{-\Psi,\tau}(y)|^2}{2\tau} - 1_\Omega(y)\Psi\big(P_{-\Psi,\tau}(y)\big).$$

It is well known that such maps are uniquely defined on $\mathcal{L}^d$-a.e. in Ω. (Indeed, $P_{\Psi,\tau}(x)$ and $P_{-\Psi,\tau}(y)$ are uniquely defined whenever the Lipschitz functions $d_{\Psi,\tau_{|\Omega}}$ and $d_{-\Psi,\tau_{|\Omega}}$ are differentiable and they are given by $P_{\Psi,\tau}(x) = x - \nabla_x d_{\Psi,\tau}$ and $P_{-\Psi,\tau}(y) = y - \nabla_y d_{-\Psi,\tau}$. Here, we are just defining them on the whole $\overline{\Omega}$ via a measurable selection argument (we omit the details).

Henceforth, $P : \overline{\Omega} \to \partial\Omega$ will be a measurable map defined in the whole $\overline{\Omega}$ with the property that

$$|x - P(x)| = d(x, \partial\Omega) \quad \forall x \in \overline{\Omega}.$$

We define the costs

$$\tilde{c}(x,y) = \frac{|x-y|^2}{2\tau} 1_{(\partial\Omega\times\partial\Omega)^c} - 1_{\partial\Omega\times\Omega}\Psi(x) + 1_{\Omega\times\partial\Omega}\Psi(y),$$

$$c(x,y) = \frac{|x-y|^2}{2\tau},$$

$$c_1 = c_{|\Omega\times\overline{\Omega}},$$

and

$$c_2 = c_{|\overline{\Omega}\times\Omega}.$$

Also, we define the set

$$\mathcal{A} = \left\{(x,y) \in \overline{\Omega}\times\overline{\Omega} \quad : \quad d_{\Psi,\tau}(x) + d_{-\Psi,\tau}(y) \geq \tilde{c}(x,y)\right\}.$$

We will work with the topological space $(\overline{\Omega}\times\overline{\Omega}, \mathcal{G})$. The topology of this space built by considering the product topology, in the spaces $\Omega\times\Omega$, $\partial\Omega\times\Omega$, $\Omega\times\partial\Omega$, and $\partial\Omega\times\partial\Omega$, and then taking the disjoint union topology. In other words, the space $\overline{\Omega}\times\overline{\Omega}$ is equipped with the topology

$$\partial\Omega\times\partial\Omega \coprod \partial\Omega\times\Omega \coprod \Omega\times\partial\Omega \coprod \Omega\times\Omega.$$

Hence, if we are given continuous functions $\{f_i\}_{i=1}^4$ from the spaces $\Omega\times\Omega$, $\partial\Omega\times\Omega, \Omega\times\partial\Omega$, and $\partial\Omega\times\partial\Omega$ to any other topological space Y, then there exists a unique continuous function $f : \overline{\Omega}\times\overline{\Omega} \to Y$ such that

$$f_i = f \circ \phi_i.$$

Here, $\{\phi_i\}_{i=1}^4$ are the canonical injections of $\Omega\times\Omega$, $\partial\Omega\times\Omega, \Omega\times\partial\Omega$, and $\partial\Omega\times\partial\Omega$ into $\overline{\Omega}\times\overline{\Omega}$. The support of the measures γ in $\overline{\Omega}\times\overline{\Omega}$ will be taken with respect to this topology. Hence, given a positive γ measure in $\overline{\Omega}\times\overline{\Omega}$, $supp(\gamma)$ is defined to be set of points (x,y) in $\overline{\Omega}\times\overline{\Omega}$ such that for every G in $\mathcal{G}$ containing (x,y), we have $\gamma(G)>0$. Additionally, we will use the notions of c-cyclical monotonicity, c-transforms, c-concavity, and c-superdifferential. We refer the reader to [5, Definitions 1.7 to 1.10]. We will only use the superdifferential. Thus, for any cost c, we will denote by $\partial^c\varphi$ the superdifferential of any c-concave function φ.

The following Proposition characterizes solutions of Problem 1.1 satisfying some hypotheses. We remark that Proposition A.2 provides conditions under which these hypotheses are satisfied.

Proposition 2.3.1. ***(Characterization of optimal pairs)*** *Let μ and ν be absolutely continuous measures in $\mathcal{M}(\Omega)$. Also, let (γ,h) be in $ADM(\mu,\nu)$. Assume that μ and $\nu+\tau h$ are strictly positive. Then, the following are equivalent:*

(i) $C_\tau(\gamma,h)$ *is minimal among all pairs in* $ADM\big(\mu,\nu\big)$ *with* h *fixed.*

(ii) γ *is concentrated on* $\mathcal{A}$ *and* $supp(\gamma)\cup\partial\Omega\times\partial\Omega$ *is* $\tilde{c}$*-cyclically monotone.*

(iii) There exist functions φ, $\varphi^*:\overline{\Omega}\to\mathbb{R}$ *having the following properties:*

(a) $\varphi_{|\Omega}$ *is* c_1-*concave,* $\varphi_{|\Omega}=(\varphi^*)^{c_1}$, $\varphi^*_{|\Omega}$ *is* c_2-*concave, and* $\varphi^*_{|\Omega}=\varphi^{c_2}$.

(b) $supp(\gamma^{\overline{\Omega}}_{\Omega})\subset\partial^{c_1}\varphi$ *and* $supp(\gamma^{\Omega}_{\overline{\Omega}})\subset\partial^{c_2}\varphi^*$.

(c) $\varphi_{|\partial\Omega}=\Psi$ *and* $\varphi^*_{|\partial\Omega}=-\Psi$.

Moreover, (γ, h) is optimal in $ADM\big(\mu, \nu\big)$ if and only if $\varphi^*_{|\Omega} = -e' \circ h + \kappa$, $\mathcal{L}^d$ a.e., for some constant κ.

Proof. We start by proving that $(i) \implies (ii)$. Define the plan $\tilde{\gamma}$ by

$$\tilde{\gamma} := \gamma_{|\mathcal{A}} + (\pi^1, P_{\Psi,\tau} \circ \pi^1)_{\#}\Big(\gamma_{|\overline{\Omega}\times\overline{\Omega}\setminus\mathcal{A}}\Big) + (P_{-\Psi,\tau} \circ \pi^2, \pi^2)_{\#}\Big(\gamma_{|\overline{\Omega}\times\overline{\Omega}\setminus\mathcal{A}}\Big).$$

Observe that $\tilde{\gamma} \in ADM(\mu, \nu)$ and

$$\begin{aligned} C_\tau(\tilde{\gamma}, h) &= \int_{\overline{\Omega}\times\overline{\Omega}} \left(\frac{|x-y|^2}{2\tau} 1_{(\partial\Omega\times\partial\Omega)^c} + \Psi(y) 1_{\Omega\times\partial\Omega} - \Psi(x) 1_{\partial\Omega\times\Omega} \right) d\tilde{\gamma} \\ &= \int_{\mathcal{A}} \tilde{c}\, d\tilde{\gamma} + \int_{\overline{\Omega}\times\overline{\Omega}\setminus\mathcal{A}} \Big(d_{\Psi,\tau}(x) + d_{-\Psi,\tau}(y) \Big)\, d\gamma \\ &\leq \int_{\overline{\Omega}\times\overline{\Omega}} \left(\frac{|x-y|^2}{2\tau} 1_{(\partial\Omega\times\partial\Omega)^c} + \Psi(y) 1_{\Omega\times\partial\Omega} - \Psi(x) 1_{\partial\Omega\times\Omega} \right) d\gamma \\ &= C_\tau(\gamma, h), \end{aligned}$$

with strict inequality if $\gamma(\overline{\Omega} \times \overline{\Omega}\setminus\mathcal{A}) > 0$. Thus, from the optimality of γ, we deduce that it is concentrated on $\mathcal{A}$.

Now we have to prove the $\tilde{c}$-cyclical monotonicity of $supp(\gamma) \cup \partial\Omega \times \partial\Omega$. Note that

$$C_\tau(\gamma, h) = C_\tau\big(\gamma + \mathcal{H}^{d-1}_{|\partial\Omega} \otimes \mathcal{H}^{d-1}_{|\partial\Omega}, h\big).$$

Hence, we can assume without loss of generality that $\partial\Omega \times \partial\Omega \subset supp(\gamma)$. Let $\{(x_i, y_i)\}_{i=1}^n \in supp(\gamma)$. Our objective is to show that

$$\sum_i \tilde{c}(x_i, y_{\sigma(i)}) - \tilde{c}(x_i, y_i) \geq 0, \quad \text{for all permutations } \sigma \text{ of } \{1, ..., n\}.$$

We proceed by contradiction. For this purpose, we assume that the above inequality fails for some permutation σ. Let

$$X_i = \begin{cases} \partial\Omega & \text{if } x_i \in \partial\Omega, \\ \Omega & \text{otherwise,} \end{cases}$$

and

$$Y_i = \begin{cases} \partial\Omega & \text{if } y_i \in \partial\Omega, \\ \Omega & \text{otherwise.} \end{cases}$$

The cost $\tilde{c}$ is continuous in $X_i \times Y_i$, for any i in $\{0, ..., n\}$. Hence, we can find neighborhoods $U_i \subset X_i$ and $V_i \subset Y_i$ of x_i and y_i such that

$$\sum_{i=1}^{N} \tilde{c}(u_i, v_{\sigma(i)}) - \tilde{c}(u_i, v_i) < 0 \quad \forall (u_i, v_i) \in U_i \times V_i \quad \text{and} \quad \forall i \in \{0, ..., n\}.$$

We will build a variation of γ, $\tilde{\gamma} = \gamma + \eta$, in such a way that its minimality is violated. To this aim, we need a signed measure η with:

(A) $\eta^- \leq \gamma$ (so that $\tilde{\gamma}$ is non-negative);

(B) $\pi^1_{\#}\eta_{|\Omega} = \pi^2_{\#}\eta_{|\Omega} = 0$ (so that $(\tilde{\gamma}, h)$ is admissible);

(C) $\int_{\overline{\Omega}\times\overline{\Omega}} \tilde{c}(x, y)\, d\eta < 0$ (so that γ is not optimal).

Let $\mathcal{C} = \Pi_{i=1}^{N} U_i \times V_i$ and $P \in \mathbb{P}(\mathcal{C})$ be defined as the product of the measures $\frac{1}{m_i}\gamma_{|U_i \times V_i}$. Here, $m_i := \gamma(U_i \times V_i)$. Denote by π^{U_i} and π^{V_i} the natural projections of $\mathcal{C}$ to U_i and V_i respectively. Also, define

$$\eta := \frac{\min_i m_i}{N} \sum_{i=1}^{N} \left(\pi^{U_i}, \pi^{V_{\sigma(i)}}\right)_{\#} P - \left(\pi^{U_i}, \pi^{V_i}\right)_{\#} P.$$

Since η satisfies $(A), (B)$, and (C), the $\tilde{c}-$cyclical monotonicity is proven. Next, we prove that $(ii) \implies (iii)$.

Arguing as Step 2 of [5, Theorem 1.13], we can produce a $\tilde{c}-concave$ function $\tilde{\varphi}$ such that $supp(\gamma) \cup \partial\Omega \times \partial\Omega \subset \partial^{\tilde{c}}\tilde{\varphi}$. Then,

$$\tilde{\varphi}(x) + \tilde{\varphi}^{\tilde{c}}(y) = \frac{|x-y|^2}{2\tau} 1_{(\partial\Omega\times\partial\Omega)^c} - \Psi(x) 1_{\partial\Omega\times\Omega} + \Psi(y) 1_{\Omega\times\partial\Omega}$$
$$\forall (x,y) \in supp(\gamma) \cup \partial\Omega \times \partial\Omega$$
$$\text{and} \quad \tilde{\varphi}(x) + \tilde{\varphi}^{\tilde{c}}(y) \leq \frac{|x-y|^2}{2\tau} 1_{(\partial\Omega\times\partial\Omega)^c} - \Psi(x) 1_{\partial\Omega\times\Omega} + \Psi(y) 1_{\Omega\times\partial\Omega}$$
$$\forall (x,y) \in \overline{\Omega} \times \overline{\Omega}. \quad (2.3.18)$$

After adding a constant, we can assume $\tilde{\varphi}^c(y_0) = 0$ for some $y_0 \in \partial\Omega$. Then, using (2.3.18) it is easy to show that $\tilde{\varphi} = 0$ on $\partial\Omega$. Consequently, $\tilde{\varphi}^c = 0$ on $\partial\Omega$ as well.

Set $\varphi = \tilde{\varphi} + 1_{\partial\Omega}\Psi$ and $\varphi^* = \tilde{\varphi}^{\tilde{c}} - \Psi 1_{\partial\Omega}$. Since the measure μ is strictly positive, by (2.3.18) we have

$$\inf_{y \in \overline{\Omega}} c(x,y) - \varphi^*(y) = \varphi(x) \quad \forall x \in \Omega.$$

Similarly, since $\pi_{2\#}\gamma$ is strictly positive, we have

$$\inf_{x \in \overline{\Omega}} c(x,y) - \varphi(x) = \varphi^*(y) \quad \forall y \in \Omega.$$

Then, all the items in (iii) can be verified using (2.3.18) (see [5, Definitions 1.7 to 1.10]).

We proceed to prove that $(iii) \implies (i)$.

Let $(\tilde{\gamma}, h)$ be any admissible pair. We set $\tilde{\varphi} = \varphi - \Psi 1_{\partial\Omega}$ and $\tilde{\varphi}^* = \varphi^* + \Psi 1_{\partial\Omega}$. By item (b) of (iii), we have that (2.3.18) holds with $\tilde{\varphi}^*$ in place of $\tilde{\varphi}^c$. Moreover, from (c) we get $\tilde{\varphi}_{|\partial\Omega} = \tilde{\varphi}^*_{|\partial\Omega} = 0$. From (a), (b), and $(B2)$, we obtain that $\tilde{\varphi}_{|\Omega}$

and $\tilde{\varphi}^*_{|\Omega}$ are Lipschitz. Thus, they are integrable against any measure with finite mass. As a consequence of these observations, we deduce

$$\begin{aligned}
C_\tau(\gamma, h) &= \int_{\overline{\Omega}\times\overline{\Omega}} \tilde{c}\, d\gamma + \tau \int_\Omega e(h)\, dx \\
&= \int_{\overline{\Omega}\times\overline{\Omega}} \Big(\tilde{\varphi}(x) + \tilde{\varphi}^*(y)\Big)\, d\gamma + \tau \int_\Omega e(h)\, dx \\
&= \int_\Omega \tilde{\varphi}(x)\, d\mu + \int_\Omega \tilde{\varphi}^*(y)\, d\nu + \tau \int_\Omega \tilde{\varphi}^*(y)\, dh + \int_\Omega e(h)\, dx \\
&= \int_{\overline{\Omega}\times\overline{\Omega}} \Big(\tilde{\varphi}(x) + \tilde{\varphi}^*(y)\Big)\, d\tilde{\gamma} + \tau \int_\Omega e(h)\, dx \\
&\leq \int_{\overline{\Omega}\times\overline{\Omega}} \tilde{c}\, d\tilde{\gamma} + \tau \int_\Omega e(h)\, dx \\
&= C_\tau(\tilde{\gamma}, h).
\end{aligned}$$

In the third and fourth line above, we have used (2.1.5). This gives us the desired implication.

To prove the last part of the Proposition, we suppose the pair (γ, h) is optimal. Also, we claim that there exists a set $L \subset \overline{\Omega}$ of zero Lebesgue measure such that for every x in $\overline{\Omega}\backslash L$ there exists $y \in \overline{\Omega}\backslash L$ such that $(x, y) \in supp(\gamma) \cup \partial\Omega \times \partial\Omega$ and

$$\begin{aligned}
e' \circ h(\tilde{y}) 1_\Omega(\tilde{y}) + \frac{|x-\tilde{y}|^2}{2\tau} 1_{(\partial\Omega\times\partial\Omega)^c}(x, \tilde{y}) + \Psi(\tilde{y}) 1_{\Omega\times\partial\Omega}(x, \tilde{y}) - \Psi(x) 1_{\partial\Omega\times\Omega}(x, \tilde{y}) \\
\geq e' \circ h(y) 1_\Omega(y) + \frac{|x-y|^2}{2\tau} 1_{(\partial\Omega\times\partial\Omega)^c}(x, y) \\
+ \Psi(y) 1_{\Omega\times\partial\Omega}(x, y) - \Psi(x) 1_{\partial\Omega\times\Omega}(x, y), \quad (2.3.19)
\end{aligned}$$

holds for every $\tilde{y}$ in $\overline{\Omega}\backslash L$. We also claim that this set L can be taken such that for every y in $\overline{\Omega}\backslash L$ there exists x in $\overline{\Omega}\backslash L$ so that $(x, y) \in supp(\gamma) \cup \partial\Omega \times \partial\Omega$ and the above inequality holds for almost every $\tilde{y}$ in $\overline{\Omega}\backslash L$. We will show these claims at the end of the proof. Now, we show how the result follows from them. Define the function

$$(-e'\circ h)^{\tilde{c}}(x) = \inf_{y\in\overline{\Omega}\backslash L} \frac{|x-y|^2}{2\tau}1_{(\partial\Omega\times\partial\Omega)^c} \\ + \Psi(y)1_{\Omega\times\partial\Omega}(x,y) - \Psi(x)1_{\partial\Omega\times\Omega}(x,y) + e'\circ h(y)1_{\Omega}(y).$$

for every x in $\overline{\Omega}\backslash L$. By (2.3.19) for every x in $\overline{\Omega}\backslash L$ there exists y in $\overline{\Omega}\backslash L$ such that $(x,y)\in supp(\gamma)\cup\partial\Omega$,

$$\begin{aligned}(-e'\circ h(y)1_{\Omega}(y)) &+ (-e'\circ h)^{\tilde{c}}(x) \\ &= \Psi(y)1_{\Omega\times\partial\Omega}(x,y) - \Psi(x)1_{\partial\Omega\times\Omega}(x,y) + \frac{|x-y|^2}{2\tau}1_{(\partial\Omega\times\partial\Omega)^c}, \\ &\qquad \text{and} \quad (-e'\circ h(y)1_{\Omega}(y)) + (-e'\circ h)^{\tilde{c}}(\tilde{x}) \\ &\qquad \le \Psi(y)1_{\Omega\times\partial\Omega}(\tilde{x},y) - \Psi(\tilde{x})1_{\partial\Omega\times\Omega}(\tilde{x},y) + \frac{|x-y|^2}{2\tau}1_{(\partial\Omega\times\partial\Omega)^c},\end{aligned}$$

for almost every $\tilde{x}$ in $\overline{\Omega}\backslash L$. Then, we have that

$$(-e'\circ h(y)1_{\Omega}(y) = \inf_{x\in\overline{\Omega}\backslash L} \Psi(y)1_{\Omega\times\partial\Omega}(x,y) \\ - \Psi(x)1_{\partial\Omega\times\Omega}(x,y) + \frac{|x-y|^2}{2\tau}1_{(\partial\Omega\times\partial\Omega)^c} - (-e'\circ h)^{\tilde{c}}(x).$$

Thus, it follows that the functions $-e'\circ h_{|\Omega\backslash L}$ admits a Lipschitz extension to Ω, which we will not relabel. Consequently, for every y in $\Omega\backslash L$ there exists $x\in\overline{\Omega}\backslash L$ such that $(x,y)\in supp(\gamma)$ and (2.3.19) holds for every $\tilde{y}\in\Omega$. Then by (2.3.19), for every y in $\Omega\backslash L$ there exists $x\in\overline{\Omega}$ and a constant $A := A(x,y)$ such that (x,y) is in $supp(\gamma)$ and

$$\tau e'\circ h(y) + \frac{|y|^2}{2} + \langle x, \tilde{y}-y\rangle + A(x,y) \le \frac{|\tilde{y}|^2}{2} + \tau e'\circ\tilde{h}(\tilde{y}), \tag{2.3.20}$$

for every $\tilde{y} \in \Omega$. Let $\mathcal{P}$ be the set of affine functions that are below $\tau e' \circ h(y) + \frac{|y|^2}{2}$ in Ω. Then, it follows that

$$\tau e' \circ h + \frac{|y|^2}{2} = \sup_{p \in \mathcal{P}} p(y),$$

for every y in $\Omega \backslash L$. This together with the Lipschitz continuity of $-e' \circ h_{|\Omega}$ implies that the function $\tau e' \circ h(y) + \frac{|y|^2}{2}$ is convex. In a similar way from (2.3.18) we can deduce that $\varphi^*_{|\Omega}$ is Lipschitz, $\frac{|y|^2}{2} - \tau\varphi^*(y)$ is convex, and for a.e y in Ω there exists a point $x \in \overline{\Omega}$ and a constant $B := B(x, y)$ such that $(x, y) \in supp(\gamma)$ and

$$-\tau\varphi^*(y) + \frac{|y|^2}{2} + \langle x, \tilde{y} - y\rangle + B(x, y) \leq \frac{|\tilde{y}|^2}{2} + \tau\varphi^*(\tilde{y}), \tag{2.3.21}$$

for every $\tilde{y} \in \Omega$. Recall $\nu + \tau h$ is absolutely continuous and uniformly bounded from below. Consequently, by Lemma A.2 $\gamma^{\Omega}_{\Omega} = (S, Id)_{\#}\nu + \tau h$, for a map S that is optimal in the classical sense and is uniquely defined a.e. Thus, it follows from (2.3.20) and (2.3.21) that $\frac{|y|^2}{2} - \tau\varphi^*(y)$ and $\tau e' \circ h(y) + \frac{|y|^2}{2}$ are Lipschitz, and have the same derivative a.e in Ω. Therefore, we deduce that there exists a constant κ such that

$$\varphi^* = -e' \circ h + \kappa \quad a.e \quad in \quad \Omega.$$

In order to prove the opposite implication, suppose $\tilde{\varphi}^* = -e' \circ h + \kappa$ and let

$(\tilde{\gamma}, \tilde{h}) \in ADM(\mu, \nu)$. When we argue as in $(iii) \implies (i)$, we obtain

$$\begin{aligned}
C_\tau(\gamma, h) &= \int_{\overline{\Omega}\times\overline{\Omega}} \tilde{c}\, d\gamma + \tau \int_\Omega e(h)\, dx \\
&= \int_{\overline{\Omega}\times\overline{\Omega}} \Big([\tilde{\varphi}(x) + \kappa] + [\tilde{\varphi}^*(y) - \kappa] \Big)\, d\gamma + \tau \int_\Omega e(h)\, dx \\
&= \int_\Omega [\tilde{\varphi}(x) + \kappa]\, d\mu + \int_\Omega [\tilde{\varphi}^*(y) - \kappa]\, d\nu + \tau \int_\Omega [\tilde{\varphi}^*(y) - \kappa] h\, dx \\
&+ \tau \int_\Omega e(h)\, dx \\
&= \int_{\overline{\Omega}\times\overline{\Omega}} \Big(\tilde{\varphi}(x) + \tilde{\varphi}^*(y) \Big)\, d\tilde{\gamma} + \tau \int_\Omega [\tilde{\varphi}^*(y) - \kappa](h - \tilde{h})\, dx \\
&+ \tau \int_\Omega e(h)\, dx \\
&\leq \int_{\overline{\Omega}\times\overline{\Omega}} \tilde{c}\, d\tilde{\gamma} + \tau \int_\Omega e(h)\, dx + \tau \int_\Omega e' \circ h(\tilde{h} - h)\, dx \\
&\leq \int_{\overline{\Omega}\times\overline{\Omega}} \tilde{c}\, d\tilde{\gamma} + \tau \int_\Omega e(\tilde{h})\, dx.
\end{aligned}$$

Here, in the last inequality we used (C2). This completes the proof of the Theorem, provided we can prove the claim.

Finally, we show (2.3.19). The idea is to use Proposition A.1 and the absolute continuity and uniform positivity of μ and $\nu + \tau h$. We only prove the statement holds for $x \in \overline{\Omega}\backslash L$; the corresponding statement for y is analogous. In order to do this we will use the same notation as in Proposition A.1.

Let L_1 be a set of zero Lebesgue measure such that every point in $\Omega\backslash L_1$ is a Lebesgue point for S, $\nu + \tau h$, and h. Also let L_2 be a set of zero Lebesgue measure such that every point in $\Omega\backslash L_2$ is a Lebesgue point for T and the density of μ. Let $A = \{y \in \Omega\backslash L_1 \,:\, S(y) \in \Omega\}$ and $B = \{x \in \Omega\backslash L_2 \,:\, T(x) \in \partial\Omega\}$. Since $\pi_{1\#}(\gamma_\Omega^\Omega + \gamma_\Omega^{\partial\Omega}) = \mu$ and $\nu + \tau h$ and μ are absolutely continuous and uniformly positive, it follows that $L_3 = \overline{\Omega}\backslash(S(A) \cup T(B))$ has zero Lebesgue measure. Set $L = L_1 \cup L_2 \cup L_3$. Then, for every $x \in \Omega\backslash L$ we have two possibilities: Either there exists $y \in \Omega\backslash L$ such that $x = S(y)$, in which cases the claim follows from (A.43) and (A.44), or $T(x) \in \partial\Omega$, in which case the claim follows from (A.45) and (A.46). It remains to consider the case when $x \in \partial\Omega\backslash L$. In such case the statement follows from (A.47) and (A.48). This concludes the proof of the Proposition. □

The following result is the analogue in our setting of Brenier's Theorem on existence and uniqueness of optimal transport maps.

Corollary 2.3.1. ***(On uniqueness of optimal pairs)*** *Let $\mu, \nu \in \mathcal{M}(\Omega)$ and fix $(\gamma, h) \in Opt(\mu, \nu)$ satisfying the hypotheses of the previous Proposition. Additionally, let φ and φ^* be the functions given by Proposition 2.3.1. Then*

(i) The function h is unique $\mathcal{L}^d$ a.e.

(ii) The plan $\gamma_{\Omega}^{\overline{\Omega}}$ is unique and it is given by $(Id, T)_{\#}\mu$. Also, $T : \Omega \to \overline{\Omega}$ is the gradient of a convex function and

$$-\nabla\varphi = \frac{T - Id}{\tau} \quad \text{a.e. in} \quad \Omega.$$

(iii) The plan $\gamma_{\overline{\Omega}}^{\Omega}$ is unique and it is given by $(S, Id)_{\#}\nu$. Also, $S : \Omega \to \overline{\Omega}$ is the gradient of a convex function and

$$-\nabla\varphi^* = \frac{S - Id}{\tau} \quad \text{a.e. in} \quad \Omega.$$

(iv) If γ has no mass concentrated on $\partial\Omega \times \partial\Omega$, then γ is unique.

Proof. By linearity of the constraint (2.1.5) in $ADM(\mu, \nu)$, the uniqueness of h follows by (C2). Due to the equivalence $(i) \iff (iii)$ of the previous Theorem, using (a) and (b) we get that the functions $\tau\varphi$ and $\tau\varphi^*$ are $\frac{d^2}{2} - concave$. Here, $d(x, y) = |x - y|$. Thus, the result follows exactly as in the classical transportation problem (see for example [6, Theorem 6.2.4 and Remark 6.2.11]). □

Henceforth we will assume, without loss of generality, that the transportation plans γ have no mass concentrated on $\partial\Omega \times \partial\Omega$.

2.4 The weak solution

In this section, we follow the minimizing movement scheme described in the introduction. This method yields a map, $t \to \rho(t)$, that belongs to $L^2_{loc}([0, \infty), W^{1,2}(\Omega))$. Such a map is a weak solution to (2.1.7). By this, we mean that the map $t \to \rho(t) - e^{\Psi - V}$ belongs to $L^2_{loc}([0, \infty), W^{1,2}_0(\Omega))$,

$$\rho(0) = \rho_0 \quad \text{in} \quad \Omega,$$

and

$$\int_\Omega \zeta\rho(s)\ dx - \int_\Omega \zeta\rho(t)\ dx = \int_t^s \bigg(\int_\Omega \big[\Delta\zeta - \langle \nabla V, \nabla\zeta\rangle\big]\rho(r)\, dx \\ - \int_\Omega \zeta[e'_x]^{-1}\big(\log(\rho(r)) + V\big)\, dx \bigg) dr,$$

for all $0 \leq t < s$ and ζ in $C_c^\infty(\Omega)$.

Similarly, we will say that a map $t \to \rho(t)$ in $L^2_{loc}([0,\infty), W^{1,2}(\Omega))$ is a weak solution of (2.1.1), if there exists a Lipschitz function $\tilde{\rho}$ such that $t \to \rho(t) - \tilde{\rho}$ belongs to $L^2_{loc}([0,\infty), W_0^{1,2}(\Omega))$,

$$\tilde{\rho} = \rho_D \quad \text{on} \quad \partial\Omega,$$

$$\rho(0) = \rho_0 \quad \text{in} \quad \Omega,$$

and

$$\int_\Omega \zeta\rho(s)\ dx - \int_\Omega \zeta\rho(t)\ dx = \int_t^s \left(\int_\Omega \left[\Delta\zeta - \langle \nabla V, \nabla\zeta \rangle\right] \rho(r)\, dx \right.$$
$$\left. - \int_\Omega \zeta F_x'(\rho(r))\, dx \right) dr,$$

for all $0 \leq t < s$ and ζ in $C_c^\infty(\Omega)$.

$$E(\mu) := \begin{cases} \int_\Omega \mathcal{E}(\rho(x), x)\, dx & if \quad \mu = \rho\, \mathcal{L}^d_{|\Omega}, \\ +\infty & otherwise, \end{cases}$$

where $\mathcal{E} : [0,\infty) \times \Omega \to [0,\infty)$ is given by

$$\mathcal{E}(z, x) := z \log z - z + V(x)z + 1.$$

We will denote by $\mathcal{E}'$ the derivative of $\mathcal{E}$ with respect to its first variable and by $D(\mathcal{E})$ the interior of the sets of points where $\mathcal{E}$ is finite. The notations $\mathcal{E}(\rho(x), x)$ and $\mathcal{E}(\rho)$ will be used interchangeably. Also, we will freely interchange $\mathcal{E}'(\rho(x), x)$ and $\mathcal{E}'(\rho)$.

The main result is the following:

Theorem 2.4.1. *With the notation from the introduction and assumptions (B1) and (B2), for any pair of functions e and Ψ satisfying (C1)-(C9), any uniformly positive and bounded initial data ρ_0, and any sequence $\tau_k \downarrow 0$ there exists a subsequence, not relabeled, such that $\rho^{\tau_k}(t)$ converges to $\rho(t)$ in $L^2(0; t_f, L^2_{loc}(\Omega))$, for any $t_f > 0$. The map $t \to \rho(t)$ belongs to $L^2_{loc}([0,\infty), W^{1,2}(\Omega))$ and is a weak solution of (2.1.7). Moreover, there exist positive constants λ and Λ such that*

$$\frac{\lambda}{\sup\{e^{-V}\}} e^{-(C_0 t + V)} \leq \rho(x,t) \leq \frac{\Lambda}{\inf\{e^{-V}\}} e^{-V}, \tag{2.4.22}$$

for almost every x.

Remark 2.4.1. *When assumptions (B1)-(B3) and (F1)-(F7) hold, and e and Ψ are as in (2.1.8) and (2.1.9), properties (C1)-(C9) hold as well and the map $t \to \rho(t)$ given by the previous Theorem is a weak solution of (2.1.1).*

The proof of Theorem 2.4.1 is involved. We begin with a technical result.

Proposition 2.4.1. ***(A step of the minimizing movement)*** *Let μ be a measure in $\mathcal{M}(\Omega)$ with the property that $E(\mu) < \infty$. Also, assume that its density is uniformly positive and bounded. Additionally, let τ be a positive number. Then, there exists a minimum $\mu_\tau \in \mathcal{M}(\Omega)$ of*

$$\rho \to E(\rho) + Wb_2^{e,\Psi,\tau}(\mu, \rho). \tag{2.4.23}$$

Moreover, there exists $\delta > 0$ such that if $\tau < \delta$, then the corresponding optimal pair $(\gamma, h) \in ADM(\mu, \mu_\tau)$ satisfies:

(i) $\mu_\tau = \rho_\tau \, \mathcal{L}^d_{|\Omega}$.

(ii) $e' \circ h = \log \rho_\tau + V$.

(iii) The restriction of γ *to* $\overline{\Omega} \times \Omega$ *is given by* $(T, Id)_{\#}\mu_\tau$. *The map* T *satisfies*

$$\frac{T(y) - y}{\tau} = \nabla \log \rho_\tau(y) + \nabla V(y), \quad \mathcal{L}^d - a.e.x. \tag{2.4.24}$$

(iv) $\rho_\tau \in W^{1,2}(\Omega)$ *and* $||Tr\,[\rho_\tau] - e^{\Psi - V}||_{L^\infty(\partial\Omega)} \leq C\sqrt{\tau}$.

Here, C *is a positive constant that depends only on* Ψ. *Also,* $Tr \,:\, W^{1,2}(\Omega) \to L^2(\partial\Omega)$ *denotes the trace operator.*

Proof. Consider a minimizing sequence of measures $\{\rho^n\}_{n=1}^{\infty}$, with corresponding optimal pairs $\{(\gamma^n, h^n)\}_{n=1}^{\infty}$ in $ADM(\mu, \rho^n)$. We claim such sequences of measures and optimal pairs have the property that the mass the elements of $\{(\rho^n, \gamma^n)\}_{n=1}^{\infty}$ and the norm in $L^1(\Omega)$ of the members of $\{h^n\}_{n=1}^{\infty}$ are uniformly bounded. Since Ω is bounded, the claim allows us to obtain compactness and produce subsequences weakly converging to γ, h, and ρ_τ . The previous convergence takes place as described in the proof of Lemma 2.3.1. We will not relabel these subsequences. The absolute continuity of h and ρ_τ is guaranteed by the superlinearity of e and $\mathcal{E}$.
The inequality

$$\liminf_{n\to\infty} E(\rho^n) \geq E(\rho_\tau),$$

is a consequence of the weak conergence, $\rho_n \rightharpoonup \rho$, and the convexity and superlinearity of the maps $\{r \to \mathcal{E}(r, x)\}_{x\in\overline{\Omega}}$ (See [6, Lemma 9.4.5], for example).

To show

$$\liminf_{n\to\infty} C_\tau(\gamma^n, h^n) \geq C_\tau(\gamma, h),$$

and $(\gamma, h) \in ADM(\mu, \rho_\tau)$, we argue as in Lemma 2.3.1. This gives us the existence of a minimum as well as item (i), assuming we can prove the claim. Next, we show the claim. Arguing as in Lemma 2.3.1 and using Jensen inequality we obtain

$$\int_\Omega \mathcal{E}(\rho)\, dx + C_\tau(\gamma, h) \geq -||\Psi||_\infty \Big(\mu(\Omega) + \nu(\Omega) + \tau|h|(\Omega)\Big)$$
$$+ K|h|(\Omega) + (C(K)-1)|\Omega| + \rho(\Omega)\log\left(\frac{\rho(\Omega)}{|\Omega|}\right) - (1 + ||V||_\infty)\rho(\Omega).$$

Taking K large enough, we obtain a uniform bound on $\rho(\Omega) + \tau|h|(\Omega)$ and consequently on $|\gamma|$, for any minimizing sequence. (Recall we assume that the plans have no mass concentrated on $\partial\Omega \times \partial\Omega$). This proves the claim.

We proceed to the proof of (ii). Let η be a function with compact support in Ω. For each $\varepsilon > 0$, let $\rho_\tau^\varepsilon = \rho_\tau - \tau\varepsilon\eta$. By Lemma 2.5.1, for sufficiently small ε we can guarantee that ρ_τ^ε is non-negative. Since $(\gamma, h + \varepsilon\eta) \in ADM(\mu, \rho_\tau^\varepsilon)$, by minimality must have

$$E(\rho_\tau^\varepsilon) - E(\rho_\tau) + C_\tau(\gamma, h + \varepsilon\eta) - C_\tau(\gamma, h) \geq 0.$$

Dividing by ε and letting $\varepsilon \downarrow 0$, due to (2.1.5), Lemma 2.5.1, Lemma A.2, the dominated convegence Theorem and the fact that e and $\mathcal{E}$ are locally Lipschitz in $D(e)$ and $D(\mathcal{E})$, we get

$$\int_\Omega \big(e' \circ h\big)\eta \; dy - \int_\Omega \big(\log \rho_\tau + V\big)\eta \; dy \geq 0.$$

Replacing η by $-\eta$ gives the desired result.

Now, we show (iii). Let λ and Λ be positive numbers such that Proposition 2.5.2 holds. Then, for $\tau \in (0,1)$ we have that ρ and ρ_τ are strictly bigger than

$$\lambda \frac{\inf e^{-V}}{\sup e^{-V}} \frac{1}{(1+C_0)}.$$

We let $\delta \in (0,1)$ have the property that Corollary 2.5.1 and Proposition A.2 hold for any $\tau \in (0,\delta)$. Now, observe that Corollary 2.3.1 and the absolute continuity of μ_τ guarantee the existence of T. Then, (iii) follows from (ii), Corollary 2.3.1, and Proposition 2.3.1 (Note that in Corollary 2.3.1, T plays the role of S).

To show (iv) we note that, by minimality of ρ_τ,

$$Wb_2^{e,\Psi,\tau}(\mu,\rho_\tau) \leq E(\mu) - E(\rho_\tau),$$

and thus

$$\begin{aligned} \frac{1}{2\tau}\int_\Omega |\nabla \log \rho_\tau + \nabla V|^2(\rho_\tau + \tau h)\,dy &\leq \frac{1}{2\tau}\int_{\overline{\Omega}\times\overline{\Omega}} |x-y|^2\,d\gamma_\tau \\ &\leq E(\mu) - E(\rho_\tau) - \tau\int_\Omega e(h)\,dy + \int_{\partial\Omega\times\Omega}\Psi(x)\,d\gamma - \int_{\Omega\times\partial\Omega}\Psi(y)\,d\gamma. \end{aligned}$$

Consequently, after makign δ smaller if necessary, we get

$$\int_\Omega |\nabla\rho_\tau|^2\,dy = C_2\int_\Omega |\nabla\log\rho_\tau|^2(\rho_\tau+\tau h)\,dy < \infty.$$

Here, $C_2 := C_2(\Psi, e, V, \rho_0)$. Also, we have used the fact that ρ_τ is bounded from below by $\lambda/(1+C_0)$, V belongs to $W^{1,2}(\Omega)$, and Corollary 2.5.1 holds.

Combining (2.5.36), Lemma 2.5.2, and Lemma 2.5.6, we can see that

$$\begin{aligned}-\frac{|y-P(y)|^2}{2\tau} - C_1|y-P(y)| - C\sqrt{\tau} &\le -\Psi(P(y)) + \log \rho_\tau(y) + V(y) \\ &\le C\sqrt{\tau} + C_1|y-P(y)| + \frac{|y-P(y)|^2}{2\tau},\end{aligned}$$

where $P(y)$ denotes any of the closest points in $\partial\Omega$ to y. Also, C and C_1 depend only on $\partial\Omega$ and Ψ. Finally (iv) follows from the previous inequality. $\square$

Proof of Theorem 2.4.1. Let ρ_0 be bounded and uniformly positive. Let $\delta \in (0,1)$ be such that Propositions 2.4.1 and 2.5.3 and Corollary 2.5.1 hold for any $\tau \in (0,\delta)$. For any n in $\mathbb{N}$, let $(\gamma_n^\tau, h_n^\tau)$ be the minimizing pair from ρ_n^τ to ρ_{n+1}^τ. Also, let T_n^τ be the map that induces $(\gamma_n^\tau)_\Omega^\Omega$ given by Proposition 2.4.1 (ii).

Let t_f be a positive number larger than τ. Iterating Proposition 2.5.2, we can see that that there exist positive constants λ and Λ such that

$$\left((1+C_0\tau)^{\frac{1}{\tau}}\right)^{-n\tau} \frac{\lambda}{\sup e^{-V}} e^{-V} \le \rho_n^\tau \le \frac{\Lambda}{\inf e^{-V}} e^{-V} \quad \forall n \in \mathbb{N}. \tag{2.4.25}$$

Note

$$\lim_{\tau\to 0}(1+C_0\tau)^{\frac{1}{\tau}} = e^{C_0}.$$

Hence, for sufficiently small τ we obtain a uniform lower bound for ρ_n^τ whenever $n\tau \le t_f + 1$. Then, Lemmas 2.5.2, 2.5.3, 2.5.4, 2.5.5, Corollary 2.5.1, and Proposition 2.5.3 can be iterated to hold, with uniform constants C, κ_1, and κ_2, for all these measures. Henceforth, we assume the condition $n\tau \le t_f + 1$. Fix $\zeta \in C_c^\infty(\Omega)$.

Recall that given γ, we denote by γ_A^B its restriction to $A \times B$. Note that since

$$\gamma_n^\tau = \big(\gamma_n^\tau\big)_\Omega^\Omega + \big(\gamma_n^\tau\big)_\Omega^{\partial\Omega} + \big(\gamma_n^\tau\big)_{\partial\Omega}^\Omega,$$

by (2.1.5) we have

$$\mu_n^\tau = \big(\pi_1\big)_\#\big(\gamma_n^\tau\big)_\Omega^\Omega + \big(\pi_1\big)_\#\big(\gamma_n^\tau\big)_\Omega^{\partial\Omega},$$

and

$$\mu_{n+1}^\tau = \big(\pi_2\big)_\#\big(\gamma_n^\tau\big)_\Omega^\Omega + \big(\pi_2\big)_\#\big(\gamma_n^\tau\big)_{\partial\Omega}^\Omega - \tau h_n^\tau \, dy.$$

Consequently, we obtain

$$\begin{aligned}\int_\Omega \zeta \, d\mu_{n+1}^\tau - \int_\Omega \zeta \, d\mu_n^\tau &= \int_{\overline{\Omega}\times\overline{\Omega}} \zeta \circ \pi_2 \, d\big(\gamma_n^\tau\big)_\Omega^\Omega - \int_{\overline{\Omega}\times\overline{\Omega}} \zeta \circ \pi_1 \, d\big(\gamma_n^\tau\big)_\Omega^\Omega \\ &\quad - \tau \int_\Omega \zeta h_n^\tau \, dy + \int_{\overline{\Omega}\times\overline{\Omega}} \zeta \circ \pi_2 \, d\big(\gamma_n^\tau\big)_{\partial\Omega}^\Omega - \int_{\overline{\Omega}\times\overline{\Omega}} \zeta \circ \pi_1 \, d\big(\gamma_n^\tau\big)_\Omega^{\partial\Omega}. \quad (2.4.26)\end{aligned}$$

First, using Proposition 2.4.1 and a Taylor expansion,

$$
\begin{aligned}
&\int_{\overline{\Omega}\times\overline{\Omega}} \zeta \circ \pi_2 \ d\big(\gamma_n^\tau\big)_\Omega^\Omega - \int_{\overline{\Omega}\times\overline{\Omega}} \zeta \circ \pi_1 \, d\big(\gamma_n^\tau\big)_\Omega^\Omega \\
&\qquad = \int_{\overline{\Omega}\times\overline{\Omega}} \big(\zeta(y) - \zeta(x)\big) \, d\big(\gamma_n^\tau\big)_\Omega^\Omega \\
&\qquad = \int_{\overline{\Omega}\times\overline{\Omega}} \big(\zeta(y) - \zeta(T_n^\tau(y))\big) 1_{\{x = T_n^\tau(y)\}} \, d\big(\gamma_n^\tau\big)_\Omega^\Omega \\
&\qquad = \int_{\overline{\Omega}\times\overline{\Omega}} (\zeta - \zeta \circ T_n^\tau) \circ \pi_2 \, d\big(\gamma_n^\tau\big)_\Omega^\Omega \\
&\qquad = \int_{\overline{\Omega}\times\overline{\Omega}} (\zeta - \zeta \circ T_n^\tau) 1_{\{T_n^\tau \notin \partial\Omega\}} \circ \pi_2 \, d\big(\gamma_n^\tau\big)_\Omega^\Omega \\
&\qquad\ + \int_{\overline{\Omega}\times\overline{\Omega}} (\zeta - \zeta \circ T_n^\tau) 1_{\{T_n^\tau \notin \partial\Omega\}} \circ \pi_2 \, d\big(\gamma_n^\tau\big)_{\partial\Omega}^\Omega \\
&\qquad = \int_\Omega (\zeta - \zeta \circ T_n^\tau) 1_{\{T_n^\tau \notin \partial\Omega\}} \, d\mu_{n+1}^\tau + R_1(\tau, n) \\
&\qquad = -\int_\Omega \langle \nabla\zeta, T_n^\tau - Id \rangle \rho_{n+1}^\tau 1_{\{T_n^\tau \notin \partial\Omega\}} \, dy + R_2(\tau, n) + R_1(\tau, n) \\
&\qquad = -\tau \int_\Omega \langle \nabla\zeta, \nabla\rho_{n+1}^\tau + \rho_{n+1}^\tau \nabla V \rangle 1_{\{T_n^\tau \notin \partial\Omega\}} \, dy + R_2(\tau, n) \\
&\qquad\ + \ R_1(\tau, n).
\end{aligned}
\tag{2.4.27}
$$

Second, by item (iii) of Proposition 2.4.1, we have

$$h_n^\tau(y) = [e'_y]^{-1}(\log \rho_{n+1}^\tau(y) + V(y))$$

and consequently

$$-\tau \int_\Omega \zeta h_n^\tau \, dy = -\tau \int_\Omega \zeta [e']^{-1}(\log \rho_{n+1}^\tau + V) \, dy.$$

Third, using Corollary 2.3.1,

$$\begin{aligned}
&\int_{\overline{\Omega}\times\overline{\Omega}} \zeta\circ\pi_2 \, d\big(\gamma_n^\tau\big)_{\partial\Omega}^{\Omega} - \int_{\overline{\Omega}\times\overline{\Omega}} \zeta\circ\pi_1 \, d\big(\gamma_n^\tau\big)_{\Omega}^{\partial\Omega} \\
&\qquad = \int_{\overline{\Omega}\times\overline{\Omega}} \zeta\circ\pi_2 1_{\{x=T_n^\tau(y)\}} \, d\big(\gamma_n^\tau\big)_{\partial\Omega}^{\Omega} - \int_{\overline{\Omega}\times\overline{\Omega}} \zeta\circ\pi_1 1_{\{S_n^\tau(x)=y\}} \, d\big(\gamma_n^\tau\big)_{\Omega}^{\partial\Omega} \\
&\qquad = \int_{\overline{\Omega}\times\overline{\Omega}} \zeta(x) 1_{\{x=T_n^\tau(y)\}} \, d\big(\gamma_n^\tau\big)_{\partial\Omega}^{\Omega} - \int_{\overline{\Omega}\times\overline{\Omega}} \zeta_{n+1}(y) 1_{\{S_n^\tau(x)=y\}} \, d\big(\gamma_n^\tau\big)_{\Omega}^{\partial\Omega} \\
&\qquad + R_3(\tau,n) \\
&\qquad = \int_{\overline{\Omega}\times\overline{\Omega}} \zeta(x) \, d((\gamma_n^\tau)_{\partial\Omega}^{\Omega} - \int_{\overline{\Omega}\times\overline{\Omega}} \zeta(y) \, d(\gamma_n^\tau)_{\Omega}^{\partial\Omega} + R_3(\tau,n).
\end{aligned}$$

Here, S_n^τ is the map which induces $(\gamma_n^\tau)_{\Omega}^{\overline{\Omega}}$, given by Corollary 2.3.1. Putting the above together, we obtain

$$\begin{aligned}
\int_\Omega \zeta \, d\mu_{n+1}^\tau - \int_\Omega \zeta \, d\mu_n^\tau = \tau\Bigg(&- \int_\Omega \langle \nabla\zeta_{n+1}, \nabla\rho_{n+1}^\tau + \rho_{n+1}^\tau \nabla V\rangle 1_{\{T_n^\tau \notin\partial\Omega\}} \, dx \\
&- \int_\Omega \zeta_{n+1} [e']^{-1}(\log\rho_{n+1}^\tau + V) \, dx \Bigg) \\
&+ \int_{\overline{\Omega}\times\overline{\Omega}} \zeta(x) \, d((\gamma_n^\tau)_{\partial\Omega}^{\Omega} - \int_{\overline{\Omega}\times\overline{\Omega}} \zeta(y) \, d(\gamma_n^\tau)_{\Omega}^{\partial\Omega} \\
&+ R(n,\tau).
\end{aligned} \tag{2.4.28}$$

Here, $R(n,\tau)$ is given by

$$
\begin{aligned}
R(n,\tau) &= R_1(n,\tau) + R_2(n,\tau) + R_3(n,\tau)\\
&= \tau \int_\Omega (\zeta(y) - \zeta \circ T_n^\tau(y)) h_n^\tau 1_{\{T_n^\tau \notin \partial\Omega\}}\, dy\\
&+ \int_\Omega \Big(\int_0^1 \Big(\langle \nabla\zeta \circ ((1-s)T_n^\tau + sId), Id - T_n^\tau \rangle\\
&- \langle \nabla\zeta, Id - T_n^\tau \rangle \Big) \rho_{n+1}^\tau 1_{\{T_n^\tau \in \partial\Omega\}}\, dy\\
&+ \int_{\overline{\Omega}\times\overline{\Omega}} \Big(\zeta \circ \pi_2 - \zeta \circ \pi_1 \Big) 1_{\{x = T_n^\tau(y)}\, d\big(\gamma_n^\tau\big)_{\partial\Omega}^{\Omega}\\
&- \int_{\overline{\Omega}\times\overline{\Omega}} \Big(\zeta \circ \pi_1 - \zeta \circ \pi_2 \Big) 1_{\{S_n^\tau(x) = y\}}\, d\big(\gamma_n^\tau\big)_{\Omega}^{\partial\Omega}.
\end{aligned}
$$

Recall ζ is compactly supported. Hence, iterating Lemma 2.5.2, for sufficiently small τ we have that the intersection between the sets $supp(\zeta\circ\pi^1)$, $supp(\zeta\circ\pi^2)$, and $supp\big(\big(\gamma_n^\tau\big)_\Omega^{\partial\Omega} + (\gamma_n^\tau\big)_{\partial\Omega}^{\Omega}\big)$ is empty. Consequently, iterating Lemmas 2.5.2 and 2.5.4, we deduce

$$
\begin{aligned}
\Big| R(n,\tau) \Big| &\leq \tau \mathrm{Lip}(\zeta) \int_\Omega |y - T_n^\tau(y)||h_n^\tau|\, dy\\
&+ \mathrm{Lip}(\nabla\zeta) \int_\Omega |T_n^\tau - Id|^2 \rho_{n+1}^\tau\, dy\\
&\leq C_1(\zeta, \Psi, e, V, \rho_0, \Omega) \Big[\tau^{\frac{3}{2}} + \int_\Omega |T_n^\tau - Id|^2 (\rho_{n+1}^\tau + \tau h)\, dy \Big].
\end{aligned}
$$

Here, we have used Corollary 2.5.1 and the fact that Ω is bounded. Now, by Proposition 2.5.3

$$
\begin{aligned}
\int_{\overline{\Omega}\times\overline{\Omega}} \frac{|x-y|^2}{2\tau}\, d\gamma_n^\tau \leq C_2(\Psi, e, V, \rho_0) \Big(E(\rho_n^\tau) - \int_\Omega \Psi\, d\mu_n^\tau\\
- E(\rho_{n+1}^\tau) + \int_\Omega \Psi\, d\mu_{n+1}^\tau + \tau \Big).
\end{aligned}
$$

Thus, combining the above inequalities with (2.1.5), Lemma 2.5.4, and Corollary 2.5.1, we get

$$\Big|R(n,\tau)\Big| \le C_3(\zeta,\Psi,e,V,\rho_0)\Bigg(\tau^{3/2} + \tau\Bigg[E(\rho_n^\tau) - \int_\Omega \Psi\, d\mu_n^\tau \\ - E(\rho_{n+1}) + \int_\Omega \Psi\, d\mu_{n+1}^\tau\Bigg]\Bigg).$$

This implies

$$\Bigg|\sum_{n=M}^{N-1} R(n,\tau)\Bigg| \le C_3(\zeta,\Psi,e,V,\rho_0)\Bigg(\sqrt{\tau}(M-N)\tau \\ + \tau\Bigg[E(\rho_M^\tau) - \int_\Omega \Psi\, d\mu_M^\tau - E(\rho_N^\tau) + \int_\Omega \Psi\, d\mu_N^\tau\Bigg]\Bigg),$$

for sufficiently small τ and all integers N and M such that $\tau M \le \tau N \le t_f + 1$.

Let $\tau = \tau_k$. Also, define

$$\rho^{\tau_k}(t) = \rho_{n+1}^{\tau_k} \quad \text{for} \quad t \in \big((n+1)\tau_k, n\tau_k\big],$$

and

$$\theta_h \rho^{\tau_k}(t) = \rho^{\tau_k}(t+h),$$

for any positive constant h. Now, choose $0 \le r < s < t_f + 1$ and add up

(2.4.28) from $M = [r\backslash\tau_k]$ to $N = [s\backslash\tau_k] - 1$ to get

$$\begin{aligned}
&\int_\Omega \zeta\rho^{\tau_k}(s)\,dx - \int_\Omega \zeta\rho^{\tau_k}(r)\,dx \\
&\qquad = \int_{\tau_k[t\backslash\tau_k]}^{\tau_k[s\backslash\tau_k]} \Big(-\int_\Omega \langle\nabla\zeta, \nabla\rho^{\tau_k}(t) + \nabla V\rho^{\tau_k}\rangle 1_{\{T_n^{\tau_k}\notin\partial\Omega;[t\backslash\tau_k]=n\}}\,dx \\
&\qquad\quad - \int_\Omega \zeta[e']^{-1}(\log\rho^{\tau_k}(t)) + V\Big)dt + \left|\sum_{n=M}^{N} R(n,\tau_k)\right| \\
&\qquad = \int_{\tau_k[t\backslash\tau_k]}^{\tau_k[s\backslash\tau_k]} \Big(\int_\Omega \big[\Delta\zeta - \langle\nabla\zeta,\nabla V\rangle\big]\rho^{\tau_k}(t)\,dx \\
&\qquad\quad - \int_\Omega \zeta[e']^{-1}(\log\rho^{\tau_k}(t) + V)\Big)dt + \sum_{n=M}^{N} R(n,\tau_k).
\end{aligned} \tag{2.4.29}$$

Here, we have used the fact that by Lemma 2.5.2, for sufficiently small τ, $\{T_n^{\tau_k} \in \partial\Omega; [t\backslash\tau_k] = n\}$ and $supp(\zeta)$ are disjoint.

The strategy to pass to the limit is to use the Aubin-Lions Theorem [91, Theorem 5]. Let U be an open set with Lipschitz boundary whose closure is compactly contained in Ω. Also, set $p > d+1$. First, note $L^2(U)$ embeds in the dual of $W^{2,p}(U)$. We will denote this space by $W^{-2,p}(U)$. Second, observe $W^{1,2}(U)$ embeds compactly in $L^2(U)$ (recall Ω is bounded). Thus, in order to use the Aubin-Lions Theorem, we will show ρ^{τ_k} is bounded in $L^2(0,t_f;L^2(U))\cap L^1_{loc}(0,t_f;W^{1,2}(U))$ and

$$||\theta_h\rho^{\tau_k} - \rho^{\tau_k}||_{L^1(t_1,t_2;W^{-2,p}(U))} \to 0 \quad \forall 0 \le t_1 < t_2 < t_f,$$

as $h \to 0$, uniformly.

Given $t \in (t_1,t_2)$ set $N = \left\lceil \frac{t+h}{\tau_k} \right\rceil - 1$ and $M = \left\lceil \frac{t}{\tau_k} \right\rceil$. For each $\zeta \in W^{2,p}(U)$, we consider an extension to $\mathbb{R}^n$ (not relabeled) satisfying $supp(\zeta) \subset \Omega$ and

$$||\zeta||_{W^{2,p}(\mathbb{R}^n)} \leq C_4 ||\zeta||_{W^{2,p}(U)}.$$

Here, $C_4 := C_4(U, \Omega)$. Then,

$$\begin{aligned}
&\int_\Omega \zeta(\theta_h \rho^{\tau_k}(t) - \rho^{\tau_k}(t))\, dx \\
&= \sum_{n=M}^{N} \int_\Omega \zeta\, d\mu_{n+1}^{\tau_k} - \int_\Omega \zeta\, d\mu_n^{\tau_k} \\
&= \sum_{n=M}^{N} \int_\Omega \zeta(y) - \zeta(x)\, d\gamma_n^{\tau_k} - \int_\Omega \zeta \tau_{\tau_k} h_n^{\tau_k}\, dx \\
&= \sum_{n=M}^{N} \int_{\overline{\Omega}\times\overline{\Omega}} \int_0^1 \langle \nabla\zeta(x + s(y-x)), y - x\rangle\, ds\, d\gamma_n^{\tau_k} - \int_\Omega \zeta \tau h_n^{\tau_k}\, dx \\
&\leq \sum_{n=M}^{N} \int_{\overline{\Omega}\times\overline{\Omega}} \left(\int_0^1 |\nabla\zeta(x+s(y-x))|^2 ds\, d\gamma_n^{\tau_k} \right)^{\frac{1}{2}} \left(\int_{\overline{\Omega}\times\overline{\Omega}} |y-x|^2\, d\gamma_n^{\tau_k} \right)^{\frac{1}{2}} \\
&\qquad\qquad + C_5 \tau_k ||\zeta||_{W^{2,p}(U)} \\
&\qquad\qquad \leq C_6 ||\zeta||_{W^{2,p}(U)} \sum_{n=M}^{N} \left[\left(\int_{\overline{\Omega}\times\overline{\Omega}} |y-x|^2\, d\gamma_n^{\tau_k} \right)^{\frac{1}{2}} + \tau_k \right].
\end{aligned}$$

Here, we used Lemma 2.5.4 and the embedding of $W^{2,p}(U)$ into $C^1(U)$. Also, $C_5 := C_5(t_f, \Psi, e, V, \rho_0)$ and $C_6 := C_6(\Omega, U, t_f, \Psi, e, V, \rho_0)$. Consequently, it follows:

$$
\begin{aligned}
&||\theta_h\rho^{\tau_k}(t)-\rho^{\tau_k}(t)||_{W^{-2,p}(U)}\\
&=\sup_{||\zeta||_{W^{2,p}(U)}=1}\int_\Omega \zeta\big(\theta_h\rho^{\tau_k}(t)-\rho^{\tau_k}(t)\big)\,dy\\
&\le C_6\sum_{n=M}^{N}\left[\left(\int_{\overline{\Omega}\times\overline{\Omega}}|y-x|^2\,d\gamma_n^\tau\right)^{\frac{1}{2}}+\tau_k\right]\\
&\le C_6\Bigg(\tau_k(N-M)+\big(\tau_k(N-M)\big)^{\frac{1}{2}}\Bigg(\sum_{n=M}^{N}\left[\left(\int_{\overline{\Omega}\times\overline{\Omega}}\frac{|y-x|^2}{\tau}\,d\gamma_n^{\tau_k}\right)\right)^{\frac{1}{2}}\\
&\le C_7\Bigg(h+\sqrt{h}\Bigg[\sum_{n=M}^{N}E(\rho_n^{\tau_k})-\int_\Omega\Psi\rho_n^{\tau_k}\,dy-E(\rho_n^{\tau_k})+\int_\Omega\Psi\rho_n^{\tau_k}+\tau\,dy\Bigg]^{\frac{1}{2}}\Bigg)\\
&\le C_7\Bigg(h+\sqrt{h}\Bigg[E(\rho_M^{\tau_k})-\int_\Omega\Psi\rho_M^{\tau_k}\,dy-E(\rho_{N+1}^{\tau_k})+\int_\Omega\Psi\rho_{N+1}^{\tau_k}+h\,dy)\Bigg]^{1/2}\Bigg).
\end{aligned}
\tag{2.4.30}
$$

Here, we used the Jensen inequality and Proposition 2.5.3. Also, $C_7 := C_7(t_f,\Omega,U,\Psi,e,V,\rho_0)$. This shows $||\theta_h\rho^{\tau_k}-\rho^{\tau_k}||_{L^1(t_1,t_2;W^{-2,p}(U))}\to 0$ as $h\to 0$, uniformly in k. In order to show that ρ^{τ_k} is bounded in $L^1(0,t_f;W^{1,2}(U))$ we use (2.1.5), Proposition 2.4.1, Lemma 2.5.4, Corollary 2.5.1 and Proposition 2.5.3 to obtain

$$
\begin{aligned}
&\int_\Omega\left|\nabla\log\rho_{n+1}^{\tau_k}+\nabla V\right|^2\big(\rho_{n+1}^{\tau_k}+\tau_k h_n^{\tau_k}\big)\,dx\,\tau_k\le\int_{\overline{\Omega}\times\overline{\Omega}}\frac{|x-y|^2}{2\tau_k}\,d\gamma_n^{\tau_k}\\
&\le C_8\Bigg(E(\rho_n^{\tau_k})-\int_\Omega\Psi\,d\mu_n^{\tau_k}-E(\rho_{n+1}^{\tau_k})+\int_\Omega\Psi\,d\mu_{n+1}^{\tau_k}+\tau_k\Bigg),
\end{aligned}
\tag{2.4.31}
$$

for every n in $[0,t_f/\tau]$. Here, $C_8:=C_8(\Psi,e,V,\rho_0)$. By Proposition (2.4.25), we have that $\tilde{\Lambda}\ge\rho_{n+1}^{\tau_k}\ge\tilde{\lambda}$, for some positive constants $\tilde{\lambda}:=\tilde{\lambda}(t_f)$ and $\tilde{\Lambda}$ and every n in $[0,t_f/\tau]$. Then, using (2.4.31), the Young inequality, Corollary

2.5.1, and the fact that V is in $W^{1,2}(\Omega)$, we get

$$\int_0^{t_f}\left(\int_\Omega |\nabla \rho^{\tau_k}(t)|^2\,dx\right)dt < C_9(t_f,\Psi,e,V,\rho_0)(1+t_f). \tag{2.4.32}$$

Hence, we conclude that $\{\rho^{\tau_k}\}_{k=1}^\infty$ is equibounded in $L^2(0,t_f;W^{1,2}(\overline{\Omega}))$. Also, from Proposition 2.5.2, we have that $\{\rho^{\tau_k}\}_{k=1}^\infty$ is equibounded in $L^2(0,t_f;L^2(\overline{\Omega}))$ as well.

This shows that the hypotheses of the Aubin-Lions Theorem are satisfied. Thus, we obtain a map $\rho \in L^2(0,t_f;L^2(U))$ and a subsequence (not relabeled). Such a subsequence satisfies that $\rho^{\tau_k} \to \rho$ in $L^2(0,t_f;L^2(U))$ as $k \to \infty$. By (2.4.30) and the Arzela Ascoli Theorem, this subsequence converges to ρ in $C^{1/2}(0,t_f;W^{-2,p}(U))$.

The final step is to use a diagonal argument along a sequence of sets U increasing to Ω. By doing this we obtain a further subsequence converging in $L^2(0,t_f;L^2_{loc}(\Omega))$ and in $C^{1/2}(0,t_f;W^{-2,p}_{loc}(\Omega))$ to a map $\rho \in L^2(0,t_f;L^2_{loc}(\Omega))$, which we have not relabeled.

Consequently, for any $\zeta \in C_c^\infty(\Omega)$,

$$\int_\Omega \zeta\rho^{\tau_k}(s)\,dx - \int_\Omega \zeta\rho^{\tau_k}(r)\,dx \to \int_\Omega \zeta\rho(s)\,dx - \int_\Omega \zeta\rho(r)\,dx.$$

Let U be an open set such that $supp(\zeta) \subset U$ and $\overline{U}$ is compactly contained in Ω. By Proposition 2.5.2, there exists $C_{10} := C_{10}(\lambda,\Lambda,t_f)$ such that

$$\int_\Omega |\zeta e'(\log\rho^{\tau_k}(t)+V)|\,dx \le C_{10}\int_\Omega ||\zeta||_{L^\infty(\Omega)}\,dx < \infty,$$

and

$$\int_\Omega |\left[\Delta\zeta - \langle\nabla\zeta,\nabla V\rangle\right]\rho^{\tau_k}(t)|\,dx \le C_{10}\int_\Omega \left[||\Delta\zeta||_{L^\infty(\Omega)} + |\nabla\zeta|^2 + |\nabla V|^2\right]dx < \infty.$$

Recall that V is in $W^{1,2}(\Omega)$. Using the fact that $\rho^{\tau_k}(t) \to \rho(t)$ in $L^2(U)$ for almost every t and the dominated convergence Theorem, we get

$$\int_\Omega \zeta e'(\log \rho^{\tau_k}(t) + V)\, dx \to \int_\Omega \zeta e'(\log \rho(t) + V)\, dx,$$

and

$$\int_\Omega \big[\Delta\zeta - \langle \nabla\zeta, \nabla V\rangle\big] \rho^{\tau_k}(t)\, dx \to \int_\Omega \big[\Delta\zeta - \langle \nabla\zeta, \nabla V\rangle\big] \rho(t)\, dx.$$

for almost every t in $[0, t_f]$. Then, a second application of the dominated convergence Theorem gives us

$$\begin{aligned}\int_{\tau_k[r\backslash\tau_k]}^{\tau_k[s\backslash\tau_k]} &\left(\int_\Omega \big[\Delta\zeta - \langle \nabla\zeta, \nabla V\rangle\big] \rho^{\tau_k}(t)\, dx - \int_\Omega \zeta [e']^{-1}(\log \rho^{\tau_k}(t) + V) \right) dt \\ &\to \int_r^s \left(\int_\Omega \big[\Delta\zeta - \langle \nabla\zeta, \nabla V\rangle\big] \rho(t)\, dx - \int_\Omega \zeta [e']^{-1}(\log \rho(t) + V) \right) dt\end{aligned}$$

Moreover, (2.4.32) and the Fatou Lemma yield

$$\int_0^{t_f} \liminf_{k\to\infty} \left(\int_\Omega |\nabla \rho^{\tau_k}(t)|^2\, dx \right) dt < \infty.$$

This gives

$$\liminf_{k\to\infty} \left(\int_\Omega |\nabla \rho^{\tau_k}(t)|^2\, dx \right) < \infty \quad \text{for a.e } t \geq 0.$$

Now, for any t such that the above $\liminf$ is finite, consider a subsequence k_n (depending on t) such that

$$\sup_{n\in\mathbb{N}} \int_\Omega |\nabla \rho^{\tau_{k_n}}(t)|^2\, dx < \infty.$$

This implies that $\rho^{\tau_{k_n}}(t)$ is uniformly bounded in $W^{1,2}(\Omega)$. Recall that $\rho^{\tau_k}(t) \to \rho(t)$ in $L^2(0, t_f; L^2_{loc}(\Omega))$. Hence, $\rho^{\tau_{k_n}}(t) \rightharpoonup \rho(t)$ in $W^{1,2}(\Omega)$. Then, by Proposition 2.4.1 we get that $\rho(t) - e^{\Psi - V}$ is in $W_0^{1,2}(\Omega)$. Hence, we have shown that the map $t \to \rho(t)$ is a weak solution of (2.1.7). Finally, to show (2.4.22), we use the fact that $\rho^{\tau_k}(t) \to \rho(t)$ in $C^{1/2}(0, t_f; W^{-2,p}_{loc}(\Omega))$ and (2.4.25).

$\square$

2.5 Properties of Minimizers

In this section, we let $\tau > 0$ be a fixed time step. We set $\mu = \rho\, \mathcal{L}^d_{|\Omega}$ and denote by $\mu_\tau = \rho_\tau\, \mathcal{L}^d_{|\Omega}$ a minimizer of (2.4.23). The density ρ is assumed to be strictly positive. Additionally, we let (γ, h) be the associated optimal pair for (ρ, ρ_τ). The objective of the section is to show some properties of μ_τ that are necessary to prove the main result, Theorem 2.4.1.

A priori, it is not immediate that one can obtain a τ independent positive lower bound for μ_τ; this is studied in Proposition 2.5.2. Consequently, we cannot use Proposition 2.3.1. However, since μ and μ_τ are absolutely continuous, Lemma A.2 guarantees the existence of maps T and S with the property that $(Id, T)_\# \mu = \gamma^{\overline{\Omega}}_{\Omega}$ and $(S, Id)_\# \mu_\tau = \gamma^{\Omega}_{\overline{\Omega}}$. We will use this maps throughout this section.
(We remark that in the proof of Proposition 4.1, we get existence of ρ_τ without using Proposition 2.3.1 or any result from this section.)

Lemma 2.5.1. ***(Boundedness and Uniform positivity)*** *The minimizer ρ_τ, defined above, is bounded and uniformly positive.*

Proof. Let r and R be positive constants such that

$$\log r + V < -\frac{\operatorname{diam}(\Omega)^2}{2\tau} - ||\Psi||_\infty,$$

$$\log R + V > \frac{\operatorname{diam}(\Omega)^2}{\tau} + 2||\Psi||_\infty,$$

and

$$R + 1 > \tau||h||_\infty.$$

(Recall that due to Lemma A.2, h is bounded). Now, set $A_R^\kappa = \{\rho_\tau > R + 1\} \cap \{\rho_\tau + \tau h < \kappa\}$ and $A_r = \{\rho_\tau + 1 < r\}$. Define $\tilde{\gamma}$ by

$$\tilde{\gamma} = \gamma + \varepsilon(P_{-\Psi,\tau,}Id)_\# 1_{A_r} - \varepsilon\gamma_{\overline{\Omega}}^{A_R^\kappa} + \varepsilon(Id, P_{\Psi,\tau,})_\# \gamma_\Omega^{A_R^\kappa}$$

If we set $\tilde{\rho} = \pi_2\tilde{\gamma} - \tau h$, then

$$\tilde{\rho} = \begin{cases} \rho_\tau & \text{in} \quad \Omega \backslash A_R^\kappa \cup A_r, \\ \rho_\tau + \varepsilon & \text{in} \quad A_r, \\ \rho_\tau - \varepsilon(\rho_\tau + \tau h) & \text{in} \quad A_R^\kappa. \end{cases}$$

Hence if $\frac{R+1}{\kappa} > \varepsilon$, then $\tilde{\rho} \in \mathcal{M}(\Omega)$ and $(\tilde{\gamma}, h) \in ADM(\mu, \tilde{\rho})$. By optimality,

$$\begin{aligned} &0 \leq E(\tilde{\rho}) - E(\rho_\tau) + C_\tau(\tilde{\gamma}, h) - C_\tau(\gamma, h) \\ &\leq \int_{A_r} \left[\mathcal{E}(\rho_\tau + \varepsilon) - \mathcal{E}(\rho_\tau)\right] dy + \varepsilon \int_{A_r} \left(\frac{|P_{-\Psi,\tau,}(y) - y|^2}{2\tau} - \Psi(P_{-\Psi,\tau,}(y)) \right) dy \\ &\quad + \int_{A_R^\kappa} \left[\mathcal{E}(\rho_\tau - \varepsilon(\rho_\tau + \tau h)) - \mathcal{E}(\rho_\tau)\right] dy + \varepsilon \int_{A_R^\kappa} \left(\frac{\operatorname{diam}^2(\Omega)}{2\tau} + ||\Psi||_\infty \right. \\ &\qquad\qquad \left. - \frac{|S(y) - y|^2}{2\tau} + ||\Psi||_\infty \right) (\rho + \tau h)\, dy. \end{aligned}$$

Then, by convexity of $\mathcal{E}$ with respect to its first variable

$$\begin{aligned} 0 \leq \varepsilon &\int_{A_r} \left[\mathcal{E}'(\rho_\tau + \varepsilon) + \frac{|P_{-\Psi,\tau,}(y) - y|^2}{2\tau} - \Psi(P_{-\Psi,\tau,}(y))\right] dy \\ &+ \varepsilon \int_{A_R^\kappa} \left[- \mathcal{E}'(\rho_\tau - \varepsilon(\rho_\tau + \tau h)) + \frac{\operatorname{diam}^2(\Omega)}{2\tau} + ||\Psi||_\infty \right. \\ &\qquad \left. - \frac{|S(y) - y|^2}{2\tau} + ||\Psi||_\infty \right] (\rho + \tau h)\, dy. \end{aligned}$$

Now if we let $\varepsilon < \min\big(1/\kappa, 1\big)$, then $\mathcal{E}'(\rho_\tau - \varepsilon(\rho_\tau + \tau h) > \log R + V$ in A_R^κ and $\mathcal{E}'(\rho_\tau + \varepsilon) < \log r + V$ in A_r. Hence, by construction both integrands are strictly negative. Thus, we conclude $|A_r| = |A_R^\kappa| = 0$. Since κ was arbitrary and $\rho_\tau + \tau h \in L^1(\Omega)$, we obtain the desired result. $\square$

In the next Proposition, we will say that a point in Ω is a density point for $\rho_\tau + \tau h$ if it is a point of with density 1 for the set $\{\rho_\tau + \tau h > 0\}$ and it is a Lebesgue point for ρ_τ and h. As before, the interior of the set of points where $\mathcal{E}$ is finite will be denoted by $D(\mathcal{E})$. The Proposition is the analogue in our context of [43, Proposition 3.7].

Proposition 2.5.1. *With the notation introduced at the beginning of this section, the following inequalities hold:*

- *Let y_1 and y_2 be points in Ω. Assume that y_1 is a density point for $\rho_\tau + \tau h$ and Lebesgue point for S and that y_2 is a Lebesgue point for ρ_τ . Then*

$$\log(\rho_\tau(y_1)) + V(y_1) + \frac{|y_1 - S(y_1)|^2}{2\tau} \leq \log(\rho_\tau(y_2)) + V(y_2) + \frac{|y_2 - S(y_1)|^2}{2\tau}. \tag{2.5.33}$$

- *Let $x \in \Omega$ be a Lebesgue point for ρ, and T and assume that $T(x) \in \partial\Omega$. Assume further that $y \in \Omega$ is a Lebesgue point for ρ_τ and h. Then*

$$\frac{|x - T(x)|^2}{2\tau} + \Psi(T(x)) \leq \log(\rho_\tau(y)) + V(y) + \frac{|x-y|^2}{2\tau}. \tag{2.5.34}$$

- *Let y_1 and $S(y_1)$ be points in Ω. Assume that y_1 is a density point for $\rho + \tau h$ and a Lebesgue point for S. Then for any $y_2 \in \partial\Omega$, we have*

$$\log(\rho_\tau(y_1)) + V(y_1) + \frac{|y_1 - S(y_1)|^2}{2\tau} \leq \frac{|y_2 - S(y_1)|^2}{2\tau} + \Psi(y_2). \quad (2.5.35)$$

- *Let $y \in \Omega$ be a density point for $\rho_\tau + \tau h$ and a Lebesgue point for $P_{-\Psi,\tau}$. Then*

$$\begin{aligned} -\frac{|y - P_{-\Psi,\tau}(y)|^2}{2\tau} &\leq -\Psi(P_{-\Psi,\tau}(y)) + \log\rho_\tau(y) + V(y) \\ &\leq \frac{|y - P_{-\Psi,\tau}(y)||y - S(y)|}{\tau} + \frac{|y - P_{-\Psi,\tau}(y)|^2}{2\tau}. \quad (2.5.36) \end{aligned}$$

- *Let $y \in \Omega$ be a density point for $\rho_\tau + \tau h$ and a Lebesgue point for S and $P_{-\Psi,\tau}$ and assume that $S(y) = P_{-\Psi,\tau}(y) \in \partial\Omega$. Then*

$$\log\big(\rho_\tau(y)\big) + V(y) + \frac{|y - P_{-\Psi,\tau}(y)|^2}{2\tau} - \Psi\big(P_{-\Psi,\tau}(y)\big) = 0. \quad (2.5.37)$$

Proof. **-Heuristic argument.** First we start with (2.5.34). Consider a point $y \in \Omega$ and suppose that we take some mass from $x \in \Omega$ and instead of sending it to $T(x) \in \partial\Omega$, we send it to y. Then, we are paying $\log\rho_\tau(y) + V(y)$ in terms of the entropy and $\frac{|x-y|^2}{2\tau}$ in terms of the cost. We are also saving $\frac{|x-T(x)|^2}{2\tau} + \Psi(T(x))$ in terms of the cost. Hence, by minimality, we must have

$$\frac{|x - T(x)|^2}{2\tau} + \Psi(T(x)) \leq \log(\rho_\tau(y)) + V(y) + \frac{|x-y|^2}{2\tau}.$$

Now we proceed with (2.5.35). We take some mass from $S(y_1)$ and instead of sending it to y_1, we send it to y_2. Then, we pay $\frac{|y_2 - S(y_1)|^2}{2\tau} + \Psi(y_2)$ in terms of the cost. We also save $\frac{|y_1 - S(y_1)|^2}{2\tau}$ in terms of the cost and $\log(\rho_\tau(y_1)) + V(y_1)$ in terms of the entropy. Thus, the desired result follows by minimality; (2.5.33) is analogous.

To show (2.5.36), we argue as follows. Pick a point $y \in \Omega$ and perturb ρ_τ by taking some small mass from a point in $P_{-\Psi,\tau}(y) \in \mathcal{P}_{-\Psi,\tau}(y)$ and putting it onto y. (In the case that $S(y) \in \partial\Omega$ we choose the point to be $S(y) = P_{-\Psi,\tau}(y) \in \mathcal{P}_{-\Psi,\tau}(y)$. It is easy to verify that the minimality of the pair allows us to do this almost everywhere in Ω.) In this way, we pay $\log(\rho_\tau(y)) + V(y)$ in terms of the entropy and $\frac{|y - P_{-\Psi,\tau}(y)|^2}{2\tau} - \Psi(P_{-\Psi,\tau}(y))$ in term of the cost. Consequently, by minimality we must have

$$\log(\rho_\tau(y)) + V(y) - \Psi(P_{-\Psi,\tau}(y)) \geq -\frac{|y - P_{-\Psi,\tau}(y)|^2}{2\tau}. \tag{2.5.38}$$

Now consider two cases. First, if $S(y) \in \Omega$, we stop sending some mass from $S(y)$ to y. Instead, we send it to $P_{-\Psi,\tau}(y) \in \partial\Omega$. By doing this, we earn $\log(\rho_\tau(y)) + V(y)$ in terms of the entropy and $\frac{|S(y)-y|^2}{2\tau}$in terms of the cost. On the other hand, we pay $\frac{|S(y) - P_{-\Psi,\tau}(y)|^2}{2\tau} + \Psi(P_{-\Psi,\tau}(y))$ in terms of the cost. Thus,

$$\begin{aligned} -\Psi(P_{-\Psi\tau}(y)) + \log(\rho_\tau(y)) + V(y) + \frac{|S(y) - (y)|^2}{2\tau} \\ \leq \frac{|S(y) - P_{-\Psi,\tau}(y)|^2}{2\tau} \\ \leq \frac{\big(|y - S(y)| + |y - P_{-\Psi,\tau}(y)|\big)^2}{2\tau}. \end{aligned}$$

Consequently, when we combine this with (2.5.38), we obtain

$$\begin{aligned}
-\frac{|y-P_{-\Psi,\tau}(y)|^2}{2\tau} &\le -\Psi(P_{-\Psi,\tau}(y)) + \log\rho_\tau(y) + V(y) \\
&\le \frac{|y-P_{-\Psi,\tau}(y)||y-S(y)|}{\tau} + \frac{|y-P_{-\Psi,\tau}(y)|^2}{2\tau}. \quad (2.5.39)
\end{aligned}$$

Second, if $S(y) \in \partial\Omega$ the above inequality is obtained as a consequence of (2.5.37) and (2.5.38).

The proof of (2.5.37) is a sort of converse of (2.5.38). Indeed, as $S(y) = P_{-\Psi,\tau}(y) \in \partial\Omega$, we know that the mass at y comes from the boundary. Hence, we can perturb ρ_τ by taking a bit less mass from the boundary, so that there is less mass in y. In this way, we save $\log(\rho_\tau(y)) + V(y)$ in terms of the entropy and $\frac{|y-P_{-\Psi,\tau}(y)|^2}{2\tau} - \Psi(P_{-\Psi,\tau}(y))$ in terms of the cost. Hence,

$$-\left(\log(\rho_\tau(y)) + V(y) + \frac{|y-P_{-\Psi,\tau}(y)|^2}{2\tau} - \Psi(P_{-\Psi,\tau}(y))\right) \ge 0.$$

From (2.5.38), we get the opposite inequality and thus we conclude the argument.

-Rigorous proof. We only prove (2.5.35); the proofs of the other inequalities are analogous.

Let $\mathcal{T}_{y_1}^{y_2} : B_r(y_1) \to \partial\Omega$ be identically equal to y_2 in $B_r(y_1)$ and let r be positive constant such that $B_r(y_1)$ is contained in Ω. Define the plan $\gamma^{r,\varepsilon}$ by

$$\gamma_r^\varepsilon = \gamma_{\overline{\Omega}}^{B_r(y_1)^c} + (1-\varepsilon)\gamma_{\overline{\Omega}}^{B_r(y_1)} + \varepsilon\big(\pi^1, \mathcal{T}_{y_1}^{y_2}\big)_\# \gamma_{\overline{\Omega}}^{B_r(y_1)}\big),$$

and set

$$\mu_\tau^{r,\varepsilon} := \pi^2_\# \gamma^{r,\varepsilon} - \tau h \, dy.$$

Observe that $\pi^1_{\#}\gamma^{r,\varepsilon} = \pi^1_{\#}\gamma$, $(\gamma^{r,\varepsilon}, h) \in ADM(\rho, \mu^{r,\varepsilon}_r)$, $(\gamma^{r,\varepsilon})^{\Omega}_{\partial\Omega} = \gamma^{\Omega}_{\partial\Omega} - \varepsilon\gamma^{B_r(y_1)}_{\partial\Omega}$, $(\gamma^{r,\varepsilon})^{\partial\Omega}_{\Omega} = \gamma^{\partial\Omega}_{\Omega} + \varepsilon\big(\pi^1, \mathcal{T}^{y_2}_{y_1}\big)_{\#}\gamma^{B_r(y_1)}_{\Omega}$, and $\mu^{r,\varepsilon}_\tau = \rho^{r,\varepsilon}_\tau \mathcal{L}^d$. Here, $\rho^{r,\varepsilon}_\tau$ is given by

$$\rho^{r,\varepsilon}_\tau = \begin{cases} \rho_\tau(y) & \text{if } y \in B_r(y_1)^c, \\ (1-\varepsilon)(\rho_\tau(y) + \tau h) - \tau h & \text{if } y \in B_r(y_1). \end{cases}$$

(We remark that by Lemma A.2 h is in $L^\infty(\Omega)$ and by Lemma 2.5.1 we that ρ_τ is bounded and uniformly positive. Hence, we can guarantee that for sufficiently small ε, $\rho^{r,\varepsilon}_\tau$ is strictly positive.)

From the minimality of ρ_τ and the relationship between γ, S, and T, we get

$$\begin{aligned}
0 &\leq \int_\Omega \mathcal{E}(\rho^{r,\varepsilon}_\tau)\,dx + C_\tau(\gamma^{r,\varepsilon}, h) - \int_\Omega \mathcal{E}(\rho_\tau)\,dx - C_\tau(\gamma, h) \\
&= \int_{B_r(y_1)} \mathcal{E}((1-\varepsilon)(\rho_\tau(y) + \tau h) - \tau h) - \mathcal{E}(\rho_\tau)\,dy \\
&+ \varepsilon \int_{B_r(y_1)} \left(\frac{|\mathcal{T}^{y_2}_{y_1}(y) - S(y)|^2}{2\tau} 1_{\{S(y)\in\Omega\}} - \frac{|y - S(y)|^2}{2\tau} \right) (\rho_\tau(y) + \tau h)\,dy \\
&+ \varepsilon \int_{B_r(y_1)} \big[\Psi(\mathcal{T}^{y_2}_{y_1}(y)) 1_{\{S(y)\in\Omega\}} + \Psi(S(y)) 1_{\{S(y)\in\partial\Omega\}}\big] (\rho_\tau(y) + \tau h)\,dy.
\end{aligned}$$

Dividing by ε and letting $\varepsilon \downarrow 0$, using Lemma 2.5.1, the dominated convergence Theorem, and the fact that $\mathcal{E}$ is Lipschitz in any compact subset of $D(\mathcal{E})$, we obtain

$$\begin{aligned}
&\int_{B_r(y_1)} \mathcal{E}'(\rho_\tau(y))(\rho_\tau(y) + \tau h)\,dy \\
&\leq \int_{B_r(y_1)} \left(\frac{|\mathcal{T}^{y_2}_{y_1}(y) - S(y)|^2}{2\tau} 1_{\{S(y)\in\Omega\}} - \frac{|y - S(y)|^2}{2\tau} \right) (\rho_\tau(y) + \tau h)\,dy \\
&\quad + \int_{B_r(y_1)} \big[\Psi(\mathcal{T}^{y_2}_{y_1}(y)) 1_{\{S(y)\in\Omega\}} + \Psi(S(y)) 1_{\{S(y)\in\partial\Omega\}}\big] (\rho_\tau(y) + \tau h)\,dy.
\end{aligned}$$

Recall that by assumption $S(y_1) \notin \partial\Omega$. Now, since y_1 is a density point for $\rho_\tau + \tau h$, and a Lebesgue point for S when we divide both sides by $\mathcal{L}^d(B_r(0))$ and we let $r \downarrow 0$, we obtain (2.5.35). $\square$

Henceforth, we will omit the proof of these kinds of perturbation arguments. They can be made rigorous using the ideas contained in the previous Proposition. In the next Proposition, the constants C_0 and s are the ones described in the introduction.

Proposition 2.5.2. *(L^∞ **Barriers**) With the notation introduced at the beginning of this section, the following holds: There exists $\varepsilon \in (0,1)$, such that if λ and Λ satisfy $0 < \lambda < \varepsilon < \frac{1}{\varepsilon} < \Lambda$ and*

$$\frac{\lambda}{\sup\{e^{-V}\}} e^{-V} \le \rho \le \frac{\Lambda}{\inf\{e^{-V}\}} e^{-V},$$

then

$$\left(\frac{1}{1 + C_0\tau}\right) \frac{\lambda}{\sup\{e^{-V}\}} e^{-V} \le \rho_\tau \le \frac{\Lambda}{\inf\{e^{-V}\}} e^{-V}.$$

Here, ε depends only on e, $||\Psi||_\infty$, and $||V||_\infty$.

Proof. We first prove the lower bound.

Assumption (C8) allows us to choose $\varepsilon \in (0, s)$ so that

$$[e'_x]^{-1}(\log r + V(x)) \le C_0\, r \quad \text{in} \quad \Omega, \tag{2.5.40}$$

and

$$\Psi > \log r + V \quad \text{on} \quad \partial\Omega,$$

for all $r \in (0, \varepsilon)$.

Let $\lambda \in (0, \varepsilon)$ such that $\lambda\, e^{-V}/\sup\{e^{-V}\} \leq \rho$ and set

$$A_\lambda = \left\{ \rho_\tau < \left(\frac{1}{1 + C_0 \tau} \right) \frac{\lambda}{\sup\{e^{-V}\}} e^{-V} \right\}.$$

For a contradiction suppose $\rho_\tau(A_\lambda) > 0$ (Note that by Lemma 2.5.1, ρ_τ is uniformly positive).

For each $x \in A_\lambda$, we perturb (γ, h) by decreasing $h(x)$ and thus increasing the mass created at x. By optimality, we get

$$\log \rho_\tau(x) + V(x) - e'(h(x)) \geq 0. \tag{2.5.41}$$

Since

$$\left(\frac{1}{1 + C_0 \tau} \right) \frac{\lambda}{\sup\{e^{-V}\}} e^{-V} < s,$$

when we combine (2.5.40) and (2.5.41) we conclude

$$h < \left(\frac{C_0}{1 + C_0 \tau} \right) \frac{\lambda}{\sup\{e^{-V}\}} e^{-V} \quad \text{in} \quad A_\lambda.$$

Let $C_\lambda = \{x \in A_\lambda \, : \, T(x) \notin A_\lambda\}$ and note that $\rho(C_\lambda) > 0$. Otherwise by (2.1.5) and the previous inequality

$$\begin{aligned}
\int_{A_\lambda} \left(\frac{1}{1 + C_0 \tau} \right) \frac{\lambda}{\sup\{e^{-V}\}} e^{-V} \, dx &> \rho_\tau(A_\lambda) \geq \rho_\tau(T(A_\lambda)) \\
&\geq \rho(T^{-1}(T(A_\lambda))) - \tau \int_{TA_\lambda} h \, dx \\
&\geq \int_{A_\lambda} \frac{\lambda}{\sup\{e^{-V}\}} e^{-V} \, dx - \left(\frac{C_0 \tau}{1 + C_0 \tau} \right) \frac{\lambda}{\sup\{e^{-V}\}} e^{-V} \, dx \\
&= \int_{A_\lambda} \left(\frac{1}{1 + C_0 \tau} \right) \frac{\lambda}{\sup\{e^{-V}\}} e^{-V} \, dx.
\end{aligned}$$

Define the sets

$$C_\lambda^1 := \left\{ x \in C_\lambda \ : \ T(x) \in \Omega \right\} \quad \text{and} \quad C_\lambda^2 := \left\{ x \in C_\lambda \ : \ T(x) \in \partial\Omega \right\}.$$

Since $C_\lambda = C_\lambda^1 \cup C_\lambda^2$, we have that either $\rho(C_\lambda^1) > 0$ or $\rho(C_\lambda^2) > 0$. Suppose we are in the first case. Then, we can find a point x which is a Lebesgue point for T such that $T(x)$ is a Lebesgue point for ρ_τ. If we stop sending some mass from x to $T(x)$, then, by optimality we obtain

$$\log \rho_\tau(x) + V(x) - \log(\rho_\tau(T(x)) - V(T(x)) - \frac{|x - T(x)|^2}{2\tau} \geq 0.$$

Since

$$\log \rho_\tau(T(x)) \geq \log \left[\left(\frac{1}{1 + C_0\tau} \right) \frac{\lambda}{\sup\{e^{-V}\}} e^{-V(T(x))} \right],$$

and

$$\log \left[\left(\frac{1}{1 + C_0\tau} \right) \frac{\lambda}{\sup\{e^{-V}\}} e^{-V(x)} \right] > \log \rho_\tau(x),$$

we get a contradiction.

Now, suppose $\rho(C_\lambda^2) > 0$. We perturb (γ, h) by not moving some mass from x to the boundary. By optimality we must have

$$\log \rho_\tau(x) + V(x) - \Psi(T(x)) - \frac{|x - T(x)|^2}{2\tau} \geq 0.$$

Since $\lambda > \rho_\tau(x)$ and $\Psi > \log \lambda + V(x)$, we get a contradiction.

Second, we prove the upper bound.

By assumptions (C1), (C7), (B1), and (B2), after making ε smaller, we can guarantee that

$$[e_x']^{-1}(\log r + V) > 0 \quad \text{in} \quad \Omega,$$

and

$$\Psi < \log r + V \quad \text{on} \quad \partial\Omega,$$

for all $r > 1/\varepsilon$.

Let $\Lambda > 1/\varepsilon$ satisfy $\frac{\Lambda}{\inf\{e^{-V}\}} e^{-V} \geq \rho$ and set $A_\Lambda = \left\{ \rho_\tau > \frac{\Lambda}{\inf\{e^{-V}\}} e^{-V} \right\}$.

In order to get a contradiction, suppose $\rho_\tau(A_\Lambda) > 0$.

For each $x \in A_\Lambda$, we perturb (γ, h) by increasing $h(x)$ and hence decreasing the amount of mass created in x. By optimality we get

$$e'(h(x), x) - \log \rho_\tau(x) - V(x) \geq 0.$$

Since $\frac{\Lambda}{\inf\{e^{-V}\}} e^{-V} \geq \Lambda$, we deduce that h is non-negative in A_Λ. Now, we consider the following cases:

Case 1: the mass of ρ_τ in A_Λ does not come from $\partial\Omega$. Let $B_\Lambda = T^{-1}(A_\Lambda)$ and observe that due to (2.1.5),

$$\int_{A_\Lambda} \frac{\Lambda}{\inf\{e^{-V}\}} e^{-V}\, dx < \rho_\tau(A_\Lambda) \leq \rho(B_\Lambda) - \tau \int_{A_\Lambda} h\, dx < \int_{B_\Lambda} \frac{\Lambda}{\inf\{e^{-V}\}} e^{-V}\, dx,$$

which implies

$$|A_\Lambda| < |B_\Lambda|.$$

Hence, we can find a Lebesgue point $x \in B_\Lambda \backslash A_\Lambda$. If we stop transporting some mass from x to $T(x)$, then by optimality, we obtain

$$-\frac{|x - T(x)|^2}{2\tau} + \log \rho_\tau(x) + V(x) - \log \rho_\tau(T(x)) - V(T(x)) \geq 0.$$

Now by construction,

$$\log \frac{\Lambda}{\sup\{e^{-V}\}} e^{-V(x)} > \log \rho_\tau(x),$$

and

$$\log \rho_\tau(T(x)) > \log \frac{\Lambda}{\sup\{e^{-V}\}} e^{-V(T(x))}.$$

When we combine this with the previous inequality we reach a contradiction.

Case 2: the mass of ρ_τ comes partially from $\partial\Omega$. Let $D_\Lambda \subset A_\Lambda$ be the set of points y such that the mass $\rho_\tau(y)$ comes from the boundary; i.e., $D_\Lambda := \{y \in A \ : \ S(y) \in \partial\Omega\}$. Also, let $y \in D_\Lambda$ be a Lebesgue point for S. Then, if we stop moving some mass from $S(y)$ to y, by optimality we obtain

$$\frac{-|S(y)-y|^2}{2\tau} + \Psi(S(y)) - \log \rho_\tau(y) - V(y) \geq 0.$$

Since $\rho_\tau(x) > \Lambda$ and $\Psi < \log \Lambda + V(x)$, we get a contradiction. This concludes the proof. $\square$

For the next Lemma we recall that we have assumed that $\gamma_{\partial\Omega}^{\partial\Omega} = 0$.

Lemma 2.5.2. ***(Transportation bound)*** *Let ε, ρ, λ, and Λ be as in Proposition 2.5.2. Then, there exists $C > 0$ such that*

$$|y - x| \leq C\sqrt{\tau} \quad \forall (x, y) \in supp(\gamma).$$

Here, C depends only on ε, λ, Λ, $||\Psi||_\infty$, and $||V||_\infty$.

Proof. Let (x, y) be a point in $supp(\gamma)$. Then, we perturb the plan γ by not moving some mass from x to y. By optimality,

$$\Psi(y)1_{\Omega\times\partial\Omega} - \Psi(x)1_{\partial\Omega\times\Omega} + \big(\log \rho_\tau(x) + V(x)\big)1_\Omega(x) \\ - \big(\log \rho_\tau(y) - V(y)\big)1_\Omega(y) - \frac{|x-y|^2}{2\tau} \geq 0.$$

Thus, the result follows (B1), (B2), (C1), and Proposition 2.5.2. $\square$

Lemma 2.5.3. ***(Boundary Mass Flux estimate)*** *Let ε, ρ, λ, and Λ be as in Proposition 2.5.2. Then, there exists $C > 0$ such that*

$$\gamma(\partial\Omega \times \Omega \cup \Omega \times \partial\Omega) \leq C\sqrt{\tau}.$$

Here, C depends only on ε, λ, Λ, $||\Psi||_\infty$, and $||V||_\infty$.

Proof. By Lemma 2.5.2, no mass either sent or taken from the boundary travels more than $C\sqrt{\tau}$. Then, at most a $C\sqrt{\tau}$ neighborhood of $\partial\Omega$ can be sent to the boundary. The mass taken from the boundary can fill at most a $C\sqrt{\tau}$ neighborhood of $\partial\Omega$. Hence, the desired result follows from (B2) and Proposition 2.5.2. □

Lemma 2.5.4. ***(Interior Mass Creation estimate)*** *Let ε, ρ, λ, and Λ be as in Proposition 2.5.2. Then, there exists $C > 0$ such that*

$$\int_\Omega |h| \; dx \leq C \quad \text{and} \quad |h| \leq C.$$

Here, C depends only on ε, Ω, λ, Λ, and $||V||_\infty$.

Proof. By item (iii) of Proposition 2.4.1, we know

$$e'(h) = \log \rho_\tau + V.$$

Consequently, by Proposition 2.5.2,

$$\log\left[\left(\frac{1}{1 + C_0\tau}\right)\lambda\right] - ||V||_\infty \leq e'_x(h(x)) \leq \log(\Lambda) + ||V||_\infty, \quad \forall x \in \Omega.$$

Using assumptions (C2), (C7), (B1) and (B2), we get that h is bounded. Thus, since Ω is bounded, the result follows. □

Corollary 2.5.1. *Let ε, ρ, λ, and Λ be as in Proposition 2.5.2. Then, there exist positive constants κ_1, κ_2, and δ such that*

$$\kappa_1 < \frac{\rho_\tau}{\rho_\tau + \tau h} < \kappa_2.$$

for every $\tau \in (0, \delta)$. Here, κ_1 and κ_2 depend only on $\varepsilon, \lambda, \Lambda, ||\Psi||_\infty$, and $||V||_\infty$.

Proof. This follows directly from Proposition 2.5.2 and Lemma 2.5.4. □

Lemma 2.5.5. ***(Boundary cost bound)*** *Let ε, ρ, λ, and Λ be as in Proposition 2.5.2. Then, for every $\epsilon > 0$, there exists $C > 0$ such that*

$$\int_{\overline{\Omega}\times\overline{\Omega}} \Psi(y)1_{\Omega\times\partial\Omega} - \Psi(x)1_{\partial\Omega\times\Omega}\, d\gamma \geq \int_\Omega \Psi\, d\mu - \int_\Omega \Psi\, d\mu_\tau \\ - \epsilon \int_{\overline{\Omega}\times\overline{\Omega}} \frac{|x-y|^2}{2\tau}\, d\gamma \; - C\left(1 + \frac{1}{\epsilon}\right)\tau.$$

Here, C depends only on $\varepsilon, \mathrm{Lip}\Psi$, λ, Λ, and $||V||_\infty$.

Proof. Set $\zeta = \Psi$ in (2.4.26). By doing this and rearranging terms, we get

$$\int_{\overline{\Omega}\times\overline{\Omega}} \Psi(y)1_{\Omega\times\partial\Omega} - \Psi(x)1_{\partial\Omega\times\Omega}\, d\gamma = \int_\Omega \Psi\, d\mu - \int_\Omega \Psi\, d\mu_\tau \\ - \left(\int_{\Omega\times\Omega} \Psi(x) - \Psi(y)\, d\gamma + \tau \int_\Omega \Psi h\, dx \right) + R(\Psi, \tau),$$

where,

$$R(\Psi, \tau) = \int_{\overline{\Omega}\times\overline{\Omega}} \left(\Psi \circ \pi_2 - \Psi \circ \pi_1\right) 1_{\{x = S(y)\}}\, d\gamma_{\partial\Omega}^{\Omega} \\ - \int_{\overline{\Omega}\times\overline{\Omega}} \left(\Psi \circ \pi_1 - \Psi \circ \pi_2\right) 1_{\{T(x) = y\}}\, d\gamma_{\Omega}^{\partial\Omega}.$$

First, by Lemma 2.5.2 and Lemma 2.5.3,

$$|R(\Psi,\tau)| \le C_1(\mathrm{Lip}\Psi, \lambda, \Lambda, \Omega, ||V||_\infty)\tau.$$

Second, by Lemma 2.5.4 we have

$$-\tau \int_\Omega \Psi h\, dx \ \ge -C_3(||\Psi||_\infty, V, \lambda, \Lambda)|\Omega|\tau.$$

Finally, by the Young inequality, Proposition 2.4.1, Proposition 2.5.2, and Lemma 2.5.4,

$$\begin{aligned} -\int_{\Omega\times\Omega} \Psi(x) - \Psi(y)\, d\gamma \ge & -\int_\Omega \mathrm{Lip}\Psi|x-y|\, d\gamma \\ & - \frac{\tau}{2\epsilon}\int (\mathrm{Lip}\Psi)^2 d\gamma - \epsilon \int_{\overline{\Omega}\times\overline{\Omega}} \frac{|x-y|^2}{2\tau}\, d\gamma \\ & \ge -C_4(\epsilon, \Psi, e, V, \lambda, \Lambda, \Omega)\frac{\tau}{\epsilon} - \epsilon \int_{\overline{\Omega}\times\overline{\Omega}} \frac{|x-y|^2}{2\tau}\, d\gamma. \end{aligned}$$

Thus, the desired result follows. □

Proposition 2.5.3. ***(Energy Inequality)*** *Let ε, ρ, λ, and Λ be as in Proposition 2.5.2. Then, there exist positive constants C and δ such that*

$$\int_{\overline{\Omega}\times\overline{\Omega}} \frac{|x-y|^2}{2\tau}\, d\gamma \le C\Big(E(\rho) - \int_\Omega \Psi\, d\mu - E(\rho_\tau) + \int_\Omega \Psi\, d\mu_\tau + \tau\Big),$$

for every $\tau \in (0,\delta)$. Here, C depends only on $\varepsilon, \lambda, \Lambda, \mathrm{Lip}\Psi$, and $||V||_\infty$.

Proof. By minimality of ρ_τ, we obtain

$$\int_{\overline{\Omega}\times\overline{\Omega}} \frac{|x-y|^2}{2\tau} + \Psi(y)1_{\Omega\times\partial\Omega} - \Psi(x)1_{\partial\Omega\times\Omega}\, d\gamma + \tau\int_\Omega e(h)\, dx + E(\rho_\tau) \le E(\rho).$$

Also, by Lemma 2.5.4 and the above inequality,

$$\int_{\overline{\Omega}\times\overline{\Omega}} \frac{|x-y|^2}{2\tau} + \Psi(y)1_{\Omega\times\partial\Omega} - \Psi(x)1_{\partial\Omega\times\Omega}\, d\gamma \leq E(\rho) - E(\rho_\tau) + C_1(\Psi, e, V, \mu, \Omega)\tau.$$

Now, using the above inequality and Lemma 2.5.5, we obtain

$$\int_{\overline{\Omega}\times\overline{\Omega}} \frac{|x-y|^2}{2\tau}\, d\gamma + \int_\Omega \Psi\, d\mu - \int_\Omega \Psi\, d\mu_\tau - \epsilon \int_{\overline{\Omega}\times\overline{\Omega}} \frac{|x-y|^2}{2\tau}\, d\gamma\, \tau - C_2\left(1+\frac{1}{\epsilon}\right)\tau$$
$$\leq E(\rho) - E(\rho_\tau) + C_1(\Psi, e, V, \lambda, \Lambda, \Omega)\tau.$$

Here, $C_2 := C_2(\Psi, e, V, \lambda, \Lambda, \Omega)$. Then, the result follows by first choosing ϵ and then δ appropriately in the above inequality. $\square$

For the next proposition, we will need the map $P\ :\rightarrow \mathbb{R}^d$, which was defined in Section 3. Such a map satisfies

$$|x - P(x)| = d(x, \partial\Omega) \quad \forall x \in \Omega.$$

Lemma 2.5.6. ***(Projection estimate)*** *Assume Ω satisfies the interior ball condition with radius $r > 0$. Then, for all x with $d(x, \partial\Omega) < \frac{r}{2}$, we have*

$$|P(x) - P_{\Psi,\tau}(x)| \leq 4\tau \mathrm{Lip}\Psi \quad \text{and} \quad |P(x) - P_{-\Psi,\tau}(x)| \leq 4\tau \mathrm{Lip}\Psi.$$

Proof. Let $x \in \Omega$ such that $d(x, \partial\Omega) < \frac{r}{2}$. By the interior ball condition, $P(x)$ is unique. For a contradiction, suppose

$$|P(x) - P_{\Psi,\tau}(x)| > 4\tau \mathrm{Lip}\Psi.$$

Denote by Q the center of the circle of radius r that is tangent to $\partial\Omega$ at $P(x)$ and is contained in Ω. Using the cosine law and the fact that $|Q - P_{\Psi,\tau}(x)| \geq r$, we can see that

$$|x - P_{\Psi,\tau}(x)|^2 - |x - P(x)|^2 \\ \geq |P(x) - P_{\Psi,\tau}(x)|^2\left(1 - \frac{|x - P(x)|}{r}\right) \geq \frac{|P(x) - P_{\Psi,\tau}(x)|^2}{2}.$$

Hence,

$$\frac{|x - P_{\Psi,\tau}(x)|^2}{2\tau} - \frac{|x - P(x)|^2}{2\tau} + \Psi(P_{\Psi,\tau}(x)) - \Psi(P(x)),$$

is bounded from below by

$$\frac{|P(x) - P_{\Psi,\tau}(x)|^2}{4\tau} - \mathrm{Lip}\Psi|P(x) - P_{\Psi,\tau}(x)|.$$

Our assumption implies that the above quantity is strictly positive. This contradicts the minimality of $P_{\Psi,\tau}(x)$. Thus, we get the first inequality of the Lemma. The second inequality can be shown using the same argument. □

Appendix

A Minimizers of problem 1.1

In this section, we study properties of the minimizers of Problem 1.1 that are needed for Section 3. For this purpose, we let μ and $\rho\, dx$ be absolutely continuous measures in $\mathcal{M}(\Omega)$ and let τ be a fixed positive number. Additionally, we define $m_r : \overline{\Omega} \to \mathbb{R}$ by

$$m_r(x) := [e'_x]^{-1}(r),$$

for any r in $\mathbb{R}$.

Henceforth, we will say that a plan is optimal in the classical sense if it is an optimal plan for the cost $d(x,y) = |x-y|^2$. Whenever γ is an optimal plan in the classical sense and $\mu = \pi_{1\#}\gamma$ is absolutely continuous, we can guarantee the existence of a map T such that $(Id, T)_{\#}\mu = \gamma$ (see, for example, [6, Theorem 6.2.4 and Remark 6.2.11]). Any map satisfying the previous property will be called optimal in the classical sense.

Lemma A.1. ***(Refinement of pairs)*** *Let μ and $\rho\, dx$ be absolutely continuous measures in $\mathcal{M}(\Omega)$ and let τ be a positive constant. Then, for any (γ, h) in $ADM(\mu, \rho)$ there exists (γ', h) and (γ'', h') in $ADM(\mu, \rho)$ with the following properties:*

(i) The plans $(\gamma')_{\Omega}^{\overline{\Omega}}$ and $(\gamma')_{\overline{\Omega}}^{\Omega}$ are optimal in the classical sense, $(\gamma')_{\partial\Omega}^{\partial\Omega} = 0$ and

$$C_\tau(\gamma', h) - C_\tau(\gamma, h) = \int_{\overline{\Omega}\times\overline{\Omega}\setminus\partial\Omega\times\partial\Omega} \frac{|x-y|^2}{2\tau}\, d\gamma' - \int_{\overline{\Omega}\times\overline{\Omega}\setminus\partial\Omega\times\partial\Omega} \frac{|x-y|^2}{2\tau}\, d\gamma.$$

(ii) We have

$$h'(x) > [e_x']^{-1}\Big(- \frac{diam(\Omega)^2}{2\tau} - ||\Psi||_\infty \Big),$$

for almost every x in Ω and

$$\begin{aligned} C_\tau(\gamma'', h') - C_\tau(\gamma', h) \le \tau \int_{A_r} \Big[e'(m_r) + \frac{|P_{-\Psi,\tau,}(y) - y|^2}{2\tau} \\ - \Psi(P_{-\Psi,\tau,}(y))\Big](m_r - h)\, dy \\ + \tau\varepsilon \int_{A_R^\kappa} \Big[- e'(h - \varepsilon(\rho + \tau h)) + \frac{diam^2(\Omega)}{2\tau} \\ - \frac{|S(y)-y|^2}{2\tau} + 2||\Psi||_\infty \Big](\rho + \tau h)\, dy \le 0. \quad \text{(A.42)} \end{aligned}$$

Here, S is an optimal map, in the classical sense, such that $(S, Id)_{\#}(\rho+\tau h) = \gamma'^{\Omega}_{\Omega}$ (this exists by the absolute continuity of $\rho + \tau h$), $A_r = \{h < m_r\}$, $A^{\kappa}_R = \{h > m_R + 1\} \cap \{\rho + \tau h < \kappa\}$, κ is a positive constant, and $\varepsilon < \min(1 \backslash \kappa, 1 \backslash \tau)$. Also, r and R are constants satisfying

$$r < -\frac{diam(\Omega)^2}{2\tau} - ||\Psi||_{\infty},$$

$$R > \frac{diam(\Omega)^2}{\tau} + 2||\Psi||_{\infty},$$

and

$$m_R > 0 \quad \text{in} \quad \Omega.$$

Proof. It is easy to verify that if $\tilde{\gamma}$ satisfies $\pi_i \gamma^{\overline{\Omega}}_{\Omega} = \pi_i (\tilde{\gamma})^{\overline{\Omega}}_{\Omega}$ and $\pi_i \gamma^{\Omega}_{\overline{\Omega}} = \pi_i (\tilde{\gamma})^{\Omega}_{\overline{\Omega}}$ for $i = 1$ and $i = 2$, then $(\tilde{\gamma}, h) \in ADM(\mu, \rho)$, and

$$C_\tau(\tilde{\gamma}, h) - C_\tau(\gamma, h) = \int_{\overline{\Omega}\times\overline{\Omega}\backslash\partial\Omega\times\partial\Omega} \frac{|x-y|^2}{2\tau}\, d\tilde{\gamma} - \int_{\overline{\Omega}\times\overline{\Omega}\backslash\partial\Omega\times\partial\Omega} \frac{|x-y|^2}{2\tau}\, d\gamma.$$

Consequently, if $(\tilde{\gamma})^{\Omega}_{\overline{\Omega}}$ and $(\tilde{\gamma})^{\overline{\Omega}}_{\Omega}$ are optimal plans in the classical sense, then

$$C_\tau(\tilde{\gamma}, h) \leq C_\tau(\gamma, h).$$

Now, (i) follows from the observation that $(\tilde{\gamma} - \tilde{\gamma}^{\partial\Omega}_{\partial\Omega}, h) \in ADM(\mu, \rho)$ and

$$C_\tau(\tilde{\gamma}, h) = C_\tau(\tilde{\gamma} - \tilde{\gamma}^{\partial\Omega}_{\partial\Omega}, h).$$

We proceed to the proof of (ii). Let γ' be the plan given by item (i). Define h' and γ'' by

$$h' = h + (m_r - h)1_{A_r} - \varepsilon \pi_2^{\#} (\gamma')^{A^{\kappa}_R}_{\overline{\Omega}},$$

$$\gamma'' = \gamma' + \tau(P_{-\Psi,\tau,}Id)_{\#}(m_r - h)1_{A_r} - \tau\varepsilon(\gamma')_{\overline{\Omega}}^{A_R^\kappa} + \tau\varepsilon(Id, P_{\Psi,\tau,})_{\#}(\gamma')_{\Omega}^{A_R^\kappa}.$$

Here, we are using same notation as in the statement of the Lemma. Observe that by (2.1.5), $(\gamma'', h') \in ADM(\mu, \rho)$,

$$h' = \begin{cases} h & \text{in} \quad \Omega \backslash A_R^\kappa \cup A_r, \\ m_r & \text{in} \quad A_r, \\ h - \varepsilon(\rho + \tau h) & \text{in} \quad A_R^\kappa, \end{cases}$$

and

$$\pi_{2\#}\gamma''_{|\Omega} = \begin{cases} \rho + \tau h & \text{in} \quad \Omega \backslash A_R^\kappa \cup A_r, \\ \rho + \tau m_r & \text{in} \quad A_r, \\ (1 - \tau\varepsilon)(\rho + \tau h) & \text{in} \quad A_R^\kappa. \end{cases}$$

Hence,

$$\begin{aligned} &C_\tau(\gamma'', h') - C_\tau(\gamma', h) \\ &\qquad \leq \tau \int_{A_r} \big[e(m_r) - e(h)\big]\, dy + \tau \int_{A_r} \bigg(\frac{|P_{-\Psi,\tau,}(y) - y|^2}{2\tau} \\ &\qquad\qquad - \Psi(P_{-\Psi,\tau,}(y)) \bigg) (m_r - h)\, dy \\ &\quad + \tau \int_{A_R^\kappa} e(h - \varepsilon(\rho + \tau h)) - e(h)\, dy + \varepsilon\tau \int_{A_R^\kappa} \bigg(\frac{\operatorname{diam}^2(\Omega)}{2\tau} + ||\Psi||_\infty \\ &\qquad\qquad - \frac{|S(y) - y|^2}{2\tau} + ||\Psi||_\infty \bigg) (\rho + \tau h)\, dy. \end{aligned}$$

Then, the desired result follows by the convexity of e with respect to its first variable and the definition of r, R, and ε. $\square$

For the next lemma, we will need the set $D(e)$, which was previously defined to be the interior of the set of points such that e is finite.

Lemma A.2. ***(Optimal maps and Bounds on the created mass)*** *Let μ and $\rho\,dx$ be absolutely continuous measures in $\mathcal{M}(\Omega)$ and let τ be a positive constant. Additionally, let (γ, h) be a pair in $Opt(\mu, \rho)$. Then*

(i) The plans $\gamma_{\Omega}^{\overline{\Omega}}$ and $\gamma_{\overline{\Omega}}^{\Omega}$ are optimal in the classical sense.

(ii) There exist maps T and S from Ω to $\overline{\Omega}$ such that

$$(Id, T)_{\#}\mu = \gamma_{\Omega}^{\overline{\Omega}},$$

and

$$(S, Id)_{\#}(\rho + \tau h) = \gamma_{\overline{\Omega}}^{\Omega}.$$

(iii) There exists a compact set $K \subset \mathbb{R} \times \overline{\Omega}$ contained in $D(e)$ such that

$$(x, h(x)) \in K,$$

for a.e x in Ω.

Proof. By Lemma A.1, (i) follows by optimality. Since μ and $\rho + \tau h$ are absolutely continuous, (ii) follows from the classical optimal transportation theory (see for example [6, Theorem 6.2.4 and Remark 6.2.11]).

Now, we proceed to the proof of (iii). Let r, R, and (γ'', h') be defined as in the previous Lemma and set $K = \{(q, x) \in \mathbb{R} \times \overline{\Omega} \,:\, m_r(x) \leq q \leq m_R(x) + 1\}$. By (C7) and (C8), K is compact. By construction, both integrands in (A.42) are strictly negative. Thus, from the minimality of (γ, h), we conclude that $|A_r| = |A_r^{\kappa}| = 0$. Since κ was arbitrary and $h + \tau\rho \in L^1(\Omega)$, we obtain the desired result. □

In the next proposition, we will say that a point in Ω is a density point for $\rho+\tau h$ if it is a point of with density 1 for the set $\{\rho+\tau h>0\}$ and it is a Lebesgue point for ρ and h.

Proposition A.1. *Let μ and $\rho\,dx$ be absolutely continuous measures in $\mathcal{M}(\Omega)$ and let τ be a positive constant. Additionally, let (γ, h) be a pair in $Opt(\mu,\rho)$. If T and S are the maps given by Lemma A.2, then the following inequalities hold:*

- *Let y_1 and y_2 be points in Ω. Assume that y_1 is a density point for $\rho+\tau h$ and a Lebesgue point for S and that y_2 is a Lebesgue point for h. Then*

$$e'(h(y_1))+\frac{|y_1-S(y_1)|^2}{2\tau}\leq e'(h(y_2))+\frac{|y_2-S(y_1)|^2}{2\tau}. \tag{A.43}$$

- *Let y_1 and $S(y_1)$ be points in Ω. Assume that y_1 is a density point for $\rho+\tau h$ and a Lebesgue point for S. Then for any $y_2\in\partial\Omega$, we have*

$$e'\circ h(y_1)+\frac{|y_1-S(y_1)|^2}{2\tau}\leq\frac{|y_2-S(y_1)|^2}{2\tau}+\Psi(y_2). \tag{A.44}$$

- *Let $x_1\in\Omega$ be a Lebesgue point for the density of μ, and T and assume that $T(x_1)\in\partial\Omega$. Assume further that $y_1\in\Omega$ is a Lebesgue point for h. Then*

$$\frac{|x_1-T(x_1)|^2}{2\tau}+\Psi(T(x_1))\leq e'(h(y_1))+\frac{|x_1-y_1|^2}{2\tau}. \tag{A.45}$$

- *Let $x_1\in\Omega$ be a Lebesgue point for the density of μ, and T and assume that $T(x_1)\in\partial\Omega$. Then for any y_1 in $\partial\Omega$,*

$$\frac{|x_1 - T(x_1)|^2}{2\tau} + \Psi(T(x_1)) \leq \frac{|x_1 - y_1|^2}{2\tau} + \Psi(y_1). \qquad \text{(A.46)}$$

- *Let $y_1 \in \Omega$ be a Lebesgue point for h. Then for any $x_1 \in \partial\Omega$,*

$$0 \leq e' \circ h(y_1) + \frac{|y_1 - x_1|^2}{2\tau} - \Psi(x_1). \qquad \text{(A.47)}$$

- *Let $y_1 \in \Omega$ be a density point for $\rho + \tau h$ such that $S(y_1) \in \partial\Omega$. Then, for any x_1 in $\partial\Omega$,*

$$\frac{|y_1 - S(y_1)|^2}{2\tau} - \Psi(S(y_1)) \leq \frac{|y_1 - x_1|^2}{2\tau} - \Psi(x_1). \qquad \text{(A.48)}$$

- *Let $y_1 \in \Omega$ be a density point for $\rho + \tau h$ such that $S(y_1) \in \partial\Omega$. Then*

$$\frac{|y_1 - S(y_1)|^2}{2\tau} - \Psi(S(y_1)) \leq -e'(h(y_1)). \qquad \text{(A.49)}$$

Proof. We only prove (A.43); the proofs of the other inequalities are analogous. Also, Proposition 2.5.1 provides heuristic arguments that illustrate the method used to prove those inequalities. This Proposition is the analogue of [43, Proposition 3.7] in our context. We have decided to include this proof since this is the first times we explain how to make these kinds of arguments rigorous with perturbations that involve mass creation.

-Heuristic argument We provide the idea to show (A.43). First suppose $S(y_1) \in \Omega$. Then we can take some mass from $S(y_1)$ and instead of sending it to y_1, we send it to y_2. In order to end up with we an admisible pair, we then have to create the missing mass at y_1 and remove the extra mass at y_2.

In order to do this, we have to decrease $h(y_1)$ and increase $h(y_2)$. By doing this we save $\frac{|y_1-S(y_1)|^2}{2\tau}$ and we pay

$$\frac{|y_2 - S(y_1)|^2}{2\tau} - e'(h_1) + e'(h_2).$$

Hence, (A.43) follows by minimality. If $S(y_1) \in \partial\Omega$, when we do the previous perturbation we save $\frac{|y_1-S(y_1)|^2}{2\tau} + \Psi(S(y_1))$ and we pay

$$\frac{|y_2 - S(y_1)|^2}{2\tau} + \Psi(S(y_1)) - e'(h_1) + e'(h_2),$$

Thus, we get the same conclusion.

-Rigorous proof We define $\gamma^{r,\varepsilon}$ and $h^{r,\varepsilon}$ by

$$\gamma_r^\varepsilon = \gamma_{\Omega}^{B_r(y_1)^c} + (1-\varepsilon)\gamma_{\Omega}^{B_r(y_1)} + \varepsilon(\pi_1, \mathcal{T}_{y_1}^{y_2})_\# \gamma_{\Omega}^{B_r(y_1)},$$

and

$$h_r^\varepsilon = h - \frac{\varepsilon}{\tau}(\pi_{2\#}\gamma_{\Omega}^{B_r(y_1)}) + \frac{\varepsilon}{\tau}(\mathcal{T}_{y_1}^{y_2} \circ \pi_{2\#}\gamma_{\Omega}^{B_r(y_1)}).$$

Here, $\mathcal{T}_{y_1}^{y_2}(y) = y - y_1 + y_2$, and r is small enough so that $B_r(y_1)$ and $B_r(y_2)$ are disjoint and contained in Ω.

Note that $\pi_\#^1(\gamma^{r,\varepsilon}) = \pi_\#^1\gamma$ and $\pi_\#^2\gamma - \tau h = \pi_\#^2\gamma^{r,\varepsilon} - \tau h^{r,\varepsilon}$. Hence, $(\gamma^{r,\varepsilon}, h^{r,\varepsilon}) \in ADM(\mu,\nu)$. By optimality, we must have

$$0 \leq C(h^{r,\varepsilon}, \gamma^{r,\varepsilon}) - C_\tau(h,\gamma).$$

Thus,

$$\begin{aligned}
0 \leq \varepsilon &\int_{B_r(y_1)} \left[\frac{|\mathcal{T}_{y_1}^{y_2} - S|^2}{2\tau} - \frac{|Id - S|^2}{2\tau}\right](\rho + \tau h)\, dy \\
&+ \tau \int_{B_r(y_1)} \left[e\left(h - \frac{\varepsilon}{\tau}(\rho_\tau + \tau h)\right) - e(h)\right] dy \\
&+ \tau \int_{B_r(y_1)} \left[e\left(h \circ \mathcal{T}_{y_1}^{y_2} + \frac{\varepsilon}{\tau}(\rho + \tau h)\right) - e(h \circ \mathcal{T}_{y_1}^{y_2})\right] dx.
\end{aligned}$$

If we divide by ε and let $\varepsilon \downarrow 0$, using Lemma A.2, the fact that e is locally Lipschitz in $D(e)$, and the dominated convergence Theorem, we obtain

$$0 \leq \int_{B_r(y_1)} \left[\frac{|\mathfrak{T}_{y_1}^{y_2} - S|^2}{2\tau} - \frac{|y-S|^2}{2\tau} \right] (\rho + \tau h)\, dy - \int_{B_r(y_1)} e'(h)(\rho + \tau h)\, dy \\ + \int_{B_r(y_1)} e'(h \circ \mathfrak{T}_{y_1}^{y_2})(\rho + \tau h)\, dy.$$

Recall y_1 is a density point for $\rho + \tau h$ and a Lebesgue point for S and y_1 and y_2 are Lebesgue points for h. Hence, when we divide by $\mathcal{L}^d(B_r(0))$ and we let $r \downarrow 0$, we obtain (A.43). $\square$

Henceforth, as we did in Section 5, we will omit the proof of these kinds of perturbation arguments. They can be made rigorous using the ideas contained in the previous Proposition. For the next proposition, we will need the sets $D(e_x)$, which were previously defined to be the interior of the set of points z such that $e(z, x)$ is finite.

Proposition A.2. ***(Bounds on the transported mass)*** *With the notations and assumptions from Proposition A.1, the following implication holds: If there there exists a positive constant λ_0 such that*

$$\lambda_0\, dx \leq \mu,$$

and

$$\lambda_0 \leq \rho,$$

then there exists a positive number δ such that

$$\frac{\lambda_0}{4} \leq \rho + \tau h,$$

for all τ in $(0,\delta)$. Here, δ depends only on λ_0, Ψ, and e.

Proof. Let $\tilde{\rho} = \rho + \tau h$. If the sets $D(e_x)$ are of the form (a, ∞) with a finite, then since $h(x) \in D(e_x)$ and (C2), (C5), and (C8) hold, the lower bound follows easily by choosing δ sufficiently small. Hence, we assume that the sets $D(e_x)$ are of the form $(-\infty, \infty)$. (We remark that due to (C8) the two conditions are mutually exclusive).

By (C6) and (C8), there exits δ such that for every $\tau < \delta$ there exists r such that

$$r < -||\Psi||_\infty$$

and

$$\frac{1}{\tau}\left(\frac{\lambda_0}{4} - \rho\right) \leq m_r \leq \frac{1}{\tau}\left(\frac{\lambda_0}{2} - \rho\right) \quad \text{in} \quad \Omega.$$

Set $A_r = \{\tilde{\rho} < \rho + \tau m_r(x)\}$ and $C_r = \{x \in A_r \ : \ T(x) \notin A_r\}$. For a contradiction, suppose $|A_r| > 0$. Note that $|C_r| \geq 0$. Otherwise, by (2.1.5)

$$\frac{\lambda_0}{2}|A_r| > \tilde{\rho}(A_r) \geq \tilde{\rho}(T(A_r)) = \mu(T^{-1}T(A_r)) \geq \lambda_0 |A_r|.$$

Define the sets

$$C_r^1 := \Big\{x \in C_r \ : \ T(x) \in \Omega\Big\} \quad \text{and} \quad C_r^2 := \Big\{x \in C_r \ : \ T(x) \in \partial\Omega\Big\}.$$

Since $C_r = C_r^1 \cup C_r^2$, we have that either $|C_r^1| > 0$ or $|C_r^2| > 0$. Suppose we are in the first case. Then, we can find a point x which is a Lebesgue point for T such that $T(x)$ is a Lebesgue point for $\tilde{\rho}$. If we stop sending some mass from x to $T(x)$ then we can create the missing mass at $T(x)$ and remove the extra mass at x. To do this we have to increase $h(x)$ and decrease $h(T(x))$. By doing this, we produce a pair in $ADM(\mu, \rho)$. Thus, by optimality, we must have

$$e'(h(x)) - e'(h(T(x)) - \frac{|x - T(x)|^2}{2\tau} \geq 0.$$

By construction, if we use (2.1.5), we obtain

$$e'(h(x)) = e'\left(\frac{\tilde{\rho}(x) - \rho(x)}{\tau}\right) < e'(m_r(x)) = r,$$

and

$$e'(h(T(x)) = e'\left(\frac{\tilde{\rho}(T(x)) - \rho(T(x))}{\tau}\right) \geq e'(m_r(x)) = r.$$

This gives us a contradiction.

Now, suppose $|C_r^2| > 0$. Then, we can find a point $x \in C_r^2$ such that x is a Lebesgue point for T and h. If we stop moving some mass from x to the boundary, then we can remove the extra mass at x. To do this we have to increase $h(x)$. By doing this, we produce a pair in $ADM(\mu, \rho)$. By optimality, we must have

$$e'(h(x)) - \Psi(T(x)) - \frac{|x - T(x)|^2}{2\tau} \geq 0.$$

As before $e'(h(x)) < r$ and by construction $r - \Psi < 0$. This gives us a contradiction. Hence, we conclude that

$$\rho + \tau h = \tilde{\rho} \geq \rho + \tau m_r \geq \frac{\lambda_0}{4} \quad \text{a.e} \quad \text{in} \quad \Omega.$$

$\square$

Chapter 3

Global well-posedness of the spatially homogeneous Kolmogorov-Vicsek model as a gradient flow

We consider the so-called spatially homogenous Kolmogorov-Vicsek model, a non-linear Fokker-Planck equation of self-driven stochastic particles with orientation interaction under the space-homogeneity. We prove the global existence and uniqueness of weak solutions to the equation. We also show that weak solutions exponentially converge to a steady state, which has the form of the Fisher-von Mises distribution.

1 Introduction

In this paper we study the dynamics of the probability density function $\rho(t,\omega)$, as one-particle distribution at time t with direction $\omega \in \mathbb{S}^{d-1}$ (unit sphere of $\mathbb{R}^d$), which satisfies the system

$$\begin{aligned} &\partial_t \rho = \Delta_\omega \rho - \nabla_\omega \cdot \Big(\rho\, \mathbb{P}_{\omega^\perp} \Omega_\rho\Big), \\ &\Omega_\rho = \frac{J_\rho}{|J_\rho|}, \quad J_\rho = \int_{\mathbb{S}^{d-1}} \omega\, \rho\, d\omega. \end{aligned} \tag{1.1}$$

Here the operators ∇_ω and Δ_ω denote the gradient and the Laplace-Beltrami operator on the sphere $\mathbb{S}^{d-1}$, respectively. The term $\mathbb{P}_{\omega^\perp}\Omega$ denotes the projection of the vector Ω onto the normal plane to ω, describing the mean-field force that governs the orientational interaction of self-driven particles by aligning them with the direction Ω determined by the flux J. Notice that Ω is not defined when $J = 0$, and this singularity in the vector field is one of the main difficulties when studying the system (1.1).

The equation (1.1) is the spatially homogeneous version of the kinetic Kolmogorov-Vicsek model, which was formally derived by Degond and Motsch [31] as a mean-field limit of the discrete Vicsek model [3, 15, 26, 50] with stochastic dynamics. Recently, the stochastic Vicsek model has received extensive attention in the mathematical topics such as the mean-field limit, hydrodynamic limit, and phase transition. Bolley, Cañizo and Carrillo [16] have rigorously justified the mean-field limit when the unit vector Ω in the force term of (1.1) is replaced by a more regular vector-field, and Degond, Frouvelle and Liu [28] provided a complete and rigorous description of phase transitions when Ω is replaced by $\nu(|J|)\Omega$, and there is a noise intensity $\tau(|J|)$ in front of $\Delta_\omega \rho$, where the functions ν and τ satisfy

$$|J| \mapsto \frac{\nu(|J|)}{|J|} \quad \text{and} \quad |J| \mapsto \tau(|J|) \quad \text{are Lipschitz and bounded.}$$

Indeed, this modification leads to the appearance of phase transitions such as the number and nature of equilibria, stability, convergence rate, phase diagram and hysteresis, which depend on the ratio between ν and τ. It is important to observe that the assumptions of ν remove the singularity of Ω because $\nu(|J|)\Omega \to 0$ as $|J| \to 0$. This phase transition problem has been studied as well in [3, 26, 28, 29, 47, 50]. Concerning studies on hydrodynamic descriptions of kinetic Vicsek model we refer to [28, 29, 31, 32, 33, 46], see also [17, 27, 52] for other related studies.

For the well-posedness of the kinetic Kolmogorov-Vicsek model, Frouvelle and Liu [47] have shown the well-posedness in the spatially homogeneous case with the "regular" force field $\mathbb{P}_{\omega^\perp} J$ instead of $\mathbb{P}_{\omega^\perp}\Omega$. Moreover they have provided the convergence rates towards equilibria by using the Onsager free energy functional and Lasalle's invariance principle, and their results have been applied in [28]. More recently, Gamba and Kang [49] proved the existence and uniqueness of weak solutions to the kinetic Kolmogorov-Vicsek model with the singular force field $\mathbb{P}_{\omega^\perp}\Omega$, under the a priori assumption that $|J| > 0$. The solution constructed in [49] are in L^∞ in time and $L^p(D)$, where D is both in

x-space and v-space. (and so they are weak solutions in the classical sense) and have stability with respect to the initial data, under the a priori assumption of $|J| > c$ uniformly in time and space. See statement of Theorem 2.1 in [49] equation (2.3) to (2.7). This statement also holds for the space-homogeneous setting under the a priori assumption of $|J| > 0$. As a study for its numerical scheme, we refer to [48].

The purpose of this paper is to present the global well-posedness and large time behavior of weak solutions to the spatially homogeneous problem (1.1). In order to prevent the singularity of Ω_ρ, we shall consider initial probability densities ρ_0 satisfying $|J_{\rho_0}| > 0$. Nonetheless, since the momentum J is not conserved, the condition $|J_{\rho_0}| > 0$ may not immediately ensure that $|J_\rho| > 0$ for all time. As we shall see, a formal computation actually does show that $|J_{\rho(t)}| \geq |J_{\rho_0}| e^{-2(d-1)t}$ (see Lemma 3.3). However, since it does not seem obvious how to justify this estimate, we shall rather argue by approximation. More precisely, we first regularize the equation (1.1) by adding a small constant $\varepsilon > 0$ to the denominator of Ω_ρ. This allows us to look at (1.1) as the gradient flow with respect to Wasserstein distance of a ε-perturbed free energy functional, and we will be able to prove the well-posedness of the regularized equation using the time-discrete scheme by Jordan, Kindeleherer, and Otto [66]. Finally, using a compactness argument, we will obtain the global well-posedness of (1.1).

For the large time behavior, we observe that, as a consequence of (2.6), the system (1.1) can be written as the nonlinear Fokker-Planck equation:

$$\partial_t \rho = \Delta_\omega \rho - \nabla_\omega \cdot \Big(\rho \, \nabla_\omega (\omega \cdot \Omega_\rho)\Big). \tag{1.2}$$

We can easily see that the equilibrium states of (1.2) have the form of the Fisher-von Mises distribution: for any given $\Omega \in \mathbb{S}^{d-1}$, these are given by

$$M_\Omega(\omega) := C_M e^{\omega \cdot \Omega},$$

where C_M is the positive constant given by

$$C_M = \frac{1}{\int_{\mathbb{S}^{d-1}} e^{\omega \cdot \Omega} \, d\omega}, \tag{1.3}$$

so that M_Ω is a probability density function. Notice that the normalization constant C_M does not depend on Ω, and can be easily computed when $d = 3$ (see Appendix). In this paper we prove that any weak solution of (1.1) converges exponentially to a stationary Fisher-von Mises distribution.

The paper is organized as follows. In the next section, we briefly present some useful results and estimates in the optimal transportation theory, and then state our main results. Section 3 is devoted to the proof of existence of weak solutions. In Section 4, we prove the convergence of weak solutions towards the equilibrium in L^1 distance. In Section 5, we show that weak solutions are locally stable with respect to the Wasserstein distance, and as a consequence we obtain the uniqueness of the weak solution.

2 Preliminaries and Main results

2.1 Probability measures on the sphere

Here we summarize useful results from optimal transportation theory that will be used throughout the paper. We consider the embedded Riemannian manifold $\mathbb{S}^{d-1} \subset \mathbb{R}^d$ endowed with the ambient metric and geodesic distance given by

$$d(x,y) := \inf\left\{ \sqrt{\int_0^1 |\, \dot{\gamma} \,|^2 \, dt} \;\Big|\; \gamma \in C^1((0,1), \mathbb{S}^{d-1}), \gamma(0) = x, \gamma(1) = y \right\}.$$

We define the 2-Wasserstein distance (or transportation distance) with quadratic cost between two probability measures μ and ν as

$$W_2(\mu,\nu) := \sqrt{\inf_{\lambda \in \Lambda(\mu,\nu)} \int_{\mathbb{S}^{d-1} \times \mathbb{S}^{d-1}} d(x,y)^2 \, d\lambda(x,y)}, \tag{2.1}$$

where $\Lambda(\mu,\nu)$ denotes the set of all probability measures λ on $\mathbb{S}^{d-1} \times \mathbb{S}^{d-1}$ with marginals μ and ν, i.e,

$$\pi_{1\#}\lambda = \mu, \quad \pi_{2\#}\gamma = \nu,$$

where $\pi_1 : (x,y) \mapsto x$ and $\pi_2 : (x,y) \mapsto y$ are the natural projections from $\mathbb{S}^{d-1} \times \mathbb{S}^{d-1}$ to $\mathbb{S}^{d-1}$, and $\pi_{1\#}\lambda$ denotes the push forward of λ through π_1. Whenever μ is absolutely continuous with respect to the volume measure of $\mathbb{S}^{d-1}$, it follows by McCann's Theorem [77] that there exists a unique optimal plan $\lambda_0 \in \Lambda(\mu,\nu)$ which minimizes (2.1), and such a plan is induced by an optimal transport map $T : \mathbb{S}^{d-1} \to \mathbb{S}^{d-1}$, i.e., $\lambda_0 = (Id, T)_{\#}\mu$ (thus, $T_{\#}\mu = \nu$). In addition, T can be written as

$$T(\omega) = \exp_\omega(\nabla\varphi(\omega)),$$

for some $d^2/2$-convex function $\varphi : \mathbb{S}^{d-1} \to \mathbb{R}$ (see for instance [5, Theorem 2.33]).

We shall denote by $(\mathcal{P}(\mathbb{S}^{d-1}), W_2)$ the metric space of probability measures on the sphere endowed with the Wasserstein distance. We recall that $(\mathcal{P}(\mathbb{S}^{d-1}), W_2)$ is a complete separable compact metric space, and a sequence μ_n converges to μ in W_2 if and only if it converges weakly in duality with functions in $C(\mathbb{S}^{d-1})$ (see for example [5, Theorem 3.7 and Remark 3.8]).

The following proposition provides a useful estimate on the directional derivative of the map $\mu \mapsto W_2^2(\mu, \nu)$, which is used in Section 3.

Proposition 2.1. *Let $\mu, \nu \in \mathcal{P}(\mathbb{S}^{d-1})$, assume that μ is absolutely continuous, let $X : \mathbb{S}^{d-1} \to T\mathbb{S}^{d-1}$ be a C^∞ vector field, and define $\mu_t := \exp(tX)_\# \mu$. Then we have*

$$\limsup_{t \to 0} \frac{W_2^2(\mu_t, \nu) - W_2^2(\mu, \nu)}{t} \leq -2 \int_{\mathbb{S}^{d-1}} \nabla_\omega \varphi(\omega) \cdot X(\omega)\, d\mu,$$

where $\varphi : \mathbb{S}^{d-1} \to \mathbb{R}$ is a $d^2/2$-convex function such that $\exp_\omega(\nabla_\omega \varphi)$ is the optimal map sending μ onto ν.

Proof. Let $\lambda_0 \in \Lambda(\mu, \nu)$ be the optimal plan, i.e., $(Id, \exp_\omega(\nabla_\omega \varphi))_\# \mu = \lambda_0$. Since the measure

$$\lambda_t := \big((\exp(tX) \circ \pi_1, \pi_2\big)_\# \lambda_0$$

belongs to $\Lambda(\mu_t, \nu)$, it follows by the definition of W_2 (see (2.1)) that

$$
\begin{aligned}
W_2^2(\mu_t, \nu) &\le \int_{\mathbb{S}^{d-1}\times\mathbb{S}^{d-1}} d(\omega, \bar{\omega})^2 \, d\lambda_t \\
&= \int_{\mathbb{S}^{d-1}\times\mathbb{S}^{d-1}} d(\exp_\omega(tX), \bar{\omega})^2 \, d\lambda_0 \\
&= \int_{\mathbb{S}^{d-1}} d\big(\exp_\omega(tX), \exp_\omega(\nabla_\omega \varphi)\big)^2 \, d\mu.
\end{aligned}
$$

We now recall that the following formula about the squared distance function (see for instance [45, Section 1.9]):

$$
\begin{aligned}
&d\big(\exp_\omega(tX(\omega)), \exp_\omega(\nabla_\omega \varphi(\omega))\big)^2 \\
&\qquad \le d\big(\omega, \exp_\omega(\nabla_\omega \varphi(\omega))\big)^2 - 2tX(\omega)\cdot\nabla_\omega\varphi(\omega) + C\,t^2. \quad (2.2)
\end{aligned}
$$

Thus

$$
\begin{aligned}
W_2^2(\mu_t, \nu) &\le \int_{\mathbb{S}^{d-1}} d\big(\omega, \exp_\omega(\nabla_\omega \varphi(\omega))\big)^2 \, d\mu - 2t \int_{\mathbb{S}^{d-1}} X(\omega)\cdot\nabla_\omega\varphi(\omega)\, d\mu + C\,t^2 \\
&= W_2^2(\mu, \nu) - 2t \int_{\mathbb{S}^{d-1}} X(\omega)\cdot\nabla_\omega\varphi(\omega)\, d\mu + C\,t^2,
\end{aligned}
$$

and the result follows. □

Throughout the paper, we mainly deal with absolutely continuous measures. Hence, by abuse of notation, we will use sometimes ρ to denote the absolutely continuous measure $\rho\, d\omega$ on the sphere $\mathbb{S}^{d-1}$.

2.2 Formulas for the calculus on the sphere

We present here some useful formulas on sphere $\mathbb{S}^{d-1}$, which are used throughout the paper.

Let $F : \mathbb{S}^{d-1} \to \mathbb{R}^d$ be a vector-valued function and $f : \mathbb{S}^{d-1} \to \mathbb{R}$ be scalar-valued function. Then we have the following formulas related to the integration by parts:

$$\int_{\mathbb{S}^{d-1}} f\,\nabla_\omega \cdot F\,d\omega = -\int_{\mathbb{S}^{d-1}} F \cdot (\nabla_\omega f - 2\omega f)\,d\omega, \tag{2.3}$$

and

$$\begin{aligned} \int_{\mathbb{S}^{d-1}} \omega\,\nabla_\omega \cdot F\,d\omega &= -\int_{\mathbb{S}^{d-1}} F\,d\omega, \\ \int_{\mathbb{S}^{d-1}} \nabla_\omega f\,d\omega &= (d-1)\int_{\mathbb{S}^{d-1}} \omega\,f\,d\omega. \end{aligned} \tag{2.4}$$

Since $\nabla_\omega f$ is a tangent vector-field, it follows immediately by the definition of the projection $\mathbb{P}_{\omega^\perp}$ that

$$\begin{aligned} &\mathbb{P}_{\omega^\perp}\omega = 0, \\ &\mathbb{P}_{\omega^\perp}\nabla_\omega f = \nabla_\omega f. \end{aligned} \tag{2.5}$$

Moreover, for any constant vector $v \in \mathbb{R}^d$ we have

$$\begin{aligned} &\nabla_\omega(\omega \cdot v) = \mathbb{P}_{\omega^\perp} v, \\ &\nabla_\omega \cdot (\mathbb{P}_{\omega^\perp} v) = -(d-1)\,\omega \cdot v. \end{aligned} \tag{2.6}$$

We refer to [47, 85] for the derivations of these formulas.

2.3 Main results

We now state our main existence, uniqueness, and convergence results. In the sequel we shall restrict to the case $d \geq 3$ since we will need to use the logarithmic Sobolev inequality on the sphere (see the proof of Lemma 4.1). We point out that Lemma 4.1 is not used in the existence part, hence our approach allows one to get existence of solutions even in the case $d = 2$.

Theorem 2.1. *(Existence and Uniqueness) Assume $d \geq 3$. Let $\rho_0 \in \mathcal{P}(\mathbb{S}^{d-1})$ be an initial probability measure satisfying*

$$|J_{\rho_0}| > 0, \quad \int_{\mathbb{S}^{d-1}} \rho_0 \log \rho_0 \, d\omega < \infty. \tag{2.7}$$

Then the equation (1.1) *has a unique weak solution $\rho \in L^2_{loc}([0,\infty), W^{1,1}(\mathbb{S}^{d-1}))$ starting from ρ_0, which is weakly continuous in time, and satisfies* (1.1) *in the weak sense: for all $\varphi \in C^\infty(\mathbb{S}^{d-1})$ and $0 \leq t < s$,*

$$\int_{\mathbb{S}^{d-1}} \varphi \left(\rho(s) - \rho(t)\right) d\omega = \int_t^s \left(\int_{\mathbb{S}^{d-1}} \left[\Delta_\omega \varphi + \nabla_\omega \varphi \cdot \nabla_\omega (\omega \cdot \Omega_{\rho(r)})\right] \rho(r)\, d\omega \right) dr.$$

Moreover for all $t > 0$,

$$|J_\rho|^2 \geq |J_{\rho_0}|^2 e^{-2(d-1)t}.$$

Theorem 2.2. *(Convergence to steady state) Assume $d \geq 3$. Let $\rho_0 \in \mathcal{P}(\mathbb{S}^{d-1})$ be an initial probability measure satisfying* (2.7)*. Then there exist a constant vector $\Omega_\infty \in \mathbb{S}^{d-1}$ and a constant $C > 0$, depending only on ρ_0 and the dimension d, such that*

$$\|\rho(t) - M_{\Omega_\infty}\|_{L^1(\mathbb{S}^{d-1})} \leq C \Big(\int_{\mathbb{S}^{d-1}} \rho_0 \log \rho_0 \, d\omega + 1 \Big) e^{-\frac{2(d-2)}{e^2} t}.$$

Remark 2.1. *Notice that, since the momentum $J_{\rho(t)}$ is not conserved in time, it is not clear how to determine the vector Ω_∞ from the initial data ρ_0.*

The following theorem provides a short time stability in Wasserstein distance when two initial probability measures are close to each other. In particular it implies uniqueness of solutions.

Theorem 2.3. *(Stability in Wasserstein distance) Assume $d \geq 3$. Let $\rho_0, \bar{\rho}_0 \in \mathcal{P}(\mathbb{S}^{d-1})$ be probability measures satisfying (2.7) and*

$$W_2(\rho_0, \bar{\rho}_0) \leq \frac{|J_{\rho_0}|}{16},$$

and let $\rho(t)$ and $\bar{\rho}(t)$ denote the solutions of (1.1) starting from ρ_0 and $\bar{\rho}_0$, respectively. Then there exist constants $C > 0$ and $\delta > 0$, depending on $\rho_0, \bar{\rho}_0$, such that

$$W_2(\rho(t), \bar{\rho}(t)) \leq e^{\lambda t}\, W_2(\rho_0, \bar{\rho}_0) \qquad \forall\, t < \delta,$$

where $\lambda := (1 + 2/|J_{\rho_0}|) - (d-2)$.

3 Existence

In this section, we prove the existence part in Theorem 2.1. For this, we first regularize the equation (1.1) using a parameter $\varepsilon \in (0,1)$ to prevent the singularity of Ω_ρ, and then take $\varepsilon \to 0$ using standard compactness argument. It is worth noticing that the existence of solutions to the regularized equation could also be proved by using the techniques in [49]. However, we used the Jordan-Kinderlehrer-Otto scheme because it gave us useful estimates for the limiting system.

3.1 Regularized equation

We first regularize (1.1) by adding $\varepsilon > 0$ to the denominator of Ω_ρ as follows:

$$\begin{aligned}
&\partial_t \rho^\varepsilon = \nabla_\omega \cdot \Big(\rho^\varepsilon \, \nabla_\omega \left(\log \rho^\varepsilon - \omega \cdot \Omega^\varepsilon_{\rho^\varepsilon}\right)\Big), \\
&\rho^\varepsilon(0) = \rho_0, \\
&\Omega^\varepsilon_{\rho^\varepsilon} = \frac{J_{\rho^\varepsilon}}{\sqrt{|J_{\rho^\varepsilon}|^2 + \varepsilon}}, \quad J_{\rho^\varepsilon} = \int_{S^{d-1}} \omega \, \rho^\varepsilon \, d\omega.
\end{aligned} \tag{3.1}$$

In the next subsections, we show the existence of weak solutions to the regularized equation (3.1) as a gradient flow with respect to Wasserstein distance of the ε-perturbed free energy functional

$$\mathcal{E}^\varepsilon(\mu) := \begin{cases} \int_{\mathbb{S}^{d-1}} \rho \log \rho \, d\omega - \sqrt{|J_\rho|^2 + \varepsilon} & \text{if } \mu = \rho \; d\omega \\ +\infty, \quad \text{otherwise.} & \end{cases}$$

Notice that since $\rho \mapsto J_\rho$ is continuous with respect to W_2, the functional $\mathcal{E}^\varepsilon$ is lower semicontinious with respect to W_2. The next lemma provides some useful properties on derivatives of the functional $\mathcal{E}^\varepsilon$.

Lemma 3.1. *For a given $\rho \in \mathcal{P}(\mathbb{S}^{d-1})$, the following results hold.*
(1) For any $\mathrm{d}^2/2$-convex function $\varphi : \mathbb{S}^{d-1} \to \mathbb{R}$, the second derivative of $\mathcal{E}^\varepsilon$ along the geodesic $\rho_t \, d\omega := \exp(t\nabla\varphi)_\# \rho \, d\omega$ at $t = 0$ is given by

$$\begin{aligned}
\frac{d^2}{dt^2}\bigg|_{t=0} \mathcal{E}^\varepsilon(\rho_t) &= \int \mathrm{tr}([D^2\varphi]^T D^2\varphi) \, \rho \, d\omega + (d-2) \int_{\mathbb{S}^{d-1}} |\nabla\varphi|^2 \rho \, d\omega \\
&\quad + \int \nabla\varphi \, D^2(\Omega^\varepsilon_\rho \cdot \omega) \, \nabla\varphi \, \rho \, d\omega \\
&- \frac{1}{\sqrt{|J_\rho|^2 + \varepsilon}} \left(\Big| \int \nabla\varphi \, \rho \, d\omega \Big|^2 - \Big(\int \Omega^\varepsilon_\rho \cdot \nabla\varphi \, \rho \, d\omega \Big)^2 \right).
\end{aligned} \tag{3.2}$$

(2) For any smooth vector field $X : \mathbb{S}^{d-1} \to T\mathbb{S}^{d-1}$, the directional derivative of $\mathcal{E}^\varepsilon$ along $\mu_t := \exp(tX)_\# \rho \, d\omega$ at $t = 0$ is given by

$$\lim_{t\to 0} \frac{\mathcal{E}^\varepsilon(\mu_t) - \mathcal{E}^\varepsilon(\rho)}{t} = \int_{S^{d-1}} \nabla_\omega (\log \rho - \omega \cdot \Omega_\rho^\varepsilon) \cdot X(\omega) \, \rho \, d\omega. \tag{3.3}$$

(3) The slope of $\mathcal{E}^\varepsilon$ is given by

$$|\nabla \mathcal{E}^\varepsilon(\rho)| := \limsup_{\bar\rho \to \rho} \frac{(\mathcal{E}^\varepsilon(\bar\rho) - \mathcal{E}^\varepsilon(\rho))_+}{W_2(\bar\rho, \rho)} = \sqrt{\int_{\mathbb{S}^{d-1}} |\nabla_\omega (\log \rho - \omega \cdot \Omega_\rho^\varepsilon)|^2 \rho \, d\omega}. \tag{3.4}$$

Proof. Once we prove (3.2), since the Hessian of the map $\omega \mapsto \Omega_\rho^\varepsilon \cdot \omega$ has norm bounded by 1 and $\mathrm{tr}([D^2\varphi]^T D^2\varphi) \geq 0$, we get

$$\left.\frac{d^2}{dt^2}\right|_{t=0} \mathcal{E}^\varepsilon(\rho_t) \geq -\lambda \int_{\mathbb{S}^{d-1}} |\nabla \varphi|^2 \rho \, d\omega, \tag{3.5}$$

with $\lambda = (1 + \varepsilon^{-1/2}) - (d-2)$. This means that the functional $\mathcal{E}^\varepsilon$ is $(-\lambda)$-convex, and it follows by standard theory (see for instance [6, Chapter 10]) that (3.3) and (3.4) hold. Thus the remaining part is devoted to the proof of (3.2).

We begin by noticing that, since $\rho_t \, d\omega := \exp(t\nabla\varphi)_\# \rho \, d\omega$ is a geodesic in W_2, the couple (ρ_t, φ_t) solves the following system of continuity/Hamilton-Jacobi equation in the distributional/viscosity sense (see for instance [94, Chapter 13]):

$$\begin{aligned} &\partial_t \rho_t + \nabla \cdot (\rho_t \nabla \varphi_t) = 0, \\ &\partial_t \varphi_t + \frac{|\nabla \varphi_t|^2}{2} = 0, \end{aligned} \tag{3.6}$$

where $\rho_0 = \rho$ and $\varphi_0 = \varphi$.

Then, using first the continuity equation above, we have

$$\begin{aligned}
\frac{d}{dt}\Big(\int_{\mathbb{S}^{d-1}} \rho_t \log \rho_t \, d\omega - \sqrt{|J_{\rho_t}|^2 + \varepsilon}\Big) &= \int \log \rho_t \, \partial_t \rho_t \, d\omega \\
&- \frac{J_{\rho_t}}{\sqrt{|J_{\rho_t}|^2 + \varepsilon}} \cdot \int \omega \, \partial_t \rho_t \, d\omega \\
&= -\int \log \rho_t \, \nabla \cdot (\rho_t \nabla \varphi_t) \, d\omega + \int \Omega^{\varepsilon}_{\rho_t} \cdot \omega \, \nabla \cdot (\rho_t \nabla \varphi_t) \, d\omega \\
&= \int \nabla \varphi_t \cdot \nabla \log \rho_t \, \rho_t \, d\omega - \int \nabla(\omega \cdot \Omega^{\varepsilon}_{\rho_t}) \cdot \nabla \varphi_t \, \rho_t \, d\omega \\
&= -\int \Delta \varphi_t \, \rho_t \, d\omega - \int \nabla(\omega \cdot \Omega^{\varepsilon}_{\rho_t}) \cdot \nabla \varphi_t \, \rho_t \, d\omega.
\end{aligned}$$

Thus,

$$\begin{aligned}
\frac{d^2}{dt^2} \mathcal{E}^{\varepsilon}(\rho_t) &= -\frac{d}{dt} \int \Delta \varphi_t \, \rho_t \, d\omega - \frac{d}{dt} \int \nabla(\omega \cdot \Omega^{\varepsilon}_{\rho_t}) \cdot \nabla \varphi_t \, \rho_t \, d\omega \\
&=: I_1 + I_2.
\end{aligned}$$

Using (3.6), we have

$$\begin{aligned}
I_1 &= \int \Delta \frac{|\nabla \varphi_t|^2}{2} \rho_t \, d\omega + \int \Delta \varphi_t \nabla \cdot (\nabla \varphi_t \, \rho_t) \, d\omega \\
&= \int \Delta \frac{|\nabla \varphi_t|^2}{2} \rho_t \, d\omega - \int \nabla \Delta \varphi_t \cdot \nabla \varphi_t \, \rho_t \, d\omega \\
&= \int \mathrm{tr}([\nabla^2 \varphi_t]^T \nabla^2 \varphi_t) \, \rho_t \, d\omega + \int \mathrm{Ric}(\nabla \varphi_t, \nabla \varphi_t) \, \rho_t \, d\omega.
\end{aligned}$$

where in the last equality we used the Bochner formula

$$\Delta \frac{|\ \nabla \varphi\ |^2}{2} - \nabla \varphi \cdot \nabla \Delta \varphi = \mathrm{tr}([\nabla^2 \varphi]^T \nabla^2 \varphi) + \mathrm{Ric}(\nabla \varphi, \nabla \varphi).$$

Since the Ricci curvature tensor of $\mathbb{S}^{d-1}$ is $(d-2) I_{d-1}$, we have

$$I_1 = \int \mathrm{tr}([\nabla^2 \varphi_t]^T \nabla^2 \varphi_t) \, \rho_t \, d\omega + (d-2) \int |\nabla \varphi_t|^2 \, \rho_t \, d\omega.$$

For I_2, we use (2.5) and (2.6) to get

$$\nabla(\omega \cdot \Omega^\varepsilon_{\rho_t}) \cdot \nabla\varphi_t = \mathbb{P}_{\omega^\perp}\Omega^\varepsilon_{\rho_t} \cdot \nabla\varphi_t = \mathbb{P}_{\omega^\perp}\nabla\varphi_t \cdot \Omega^\varepsilon_{\rho_t} = \nabla\varphi_t \cdot \Omega^\varepsilon_{\rho_t}, \tag{3.7}$$

which yields

$$\begin{aligned} I_2 &= -\frac{d}{dt}\int \nabla\varphi_t \cdot \Omega^\varepsilon_{\rho_t}\,\rho_t\,d\omega \\ &= -\int \partial_t\rho_t\,\nabla\varphi_t \cdot \Omega^\varepsilon_{\rho_t}\,d\omega - \int \rho_t\,\nabla\partial_t\varphi_t \cdot \Omega^\varepsilon_{\rho_t}\,d\omega - \int \rho_t\,\nabla\varphi_t \cdot \partial_t\Omega^\varepsilon_{\rho_t}\,d\omega \\ &=: I_{21} + I_{22} + I_{23}. \end{aligned}$$

Using (3.6) and (3.7), we have

$$\begin{aligned} I_{21} &= \int \nabla \cdot (\rho_t\nabla\varphi_t)\,\nabla\varphi_t \cdot \Omega^\varepsilon_{\rho_t}\,d\omega \\ &= \int \nabla \cdot (\rho_t\nabla\varphi_t)\,\nabla(\omega \cdot \Omega^\varepsilon_{\rho_t}) \cdot \nabla\varphi_t\,d\omega \\ &= -\int \rho_t\,\nabla\varphi_t \cdot \nabla\Big(\nabla(\omega \cdot \Omega^\varepsilon_{\rho_t}) \cdot \nabla\varphi_t\Big)\,d\omega \\ &= -\int \rho_t\,\nabla\varphi_t\,D^2(\omega \cdot \Omega^\varepsilon_{\rho_t})\,\nabla\varphi_t\,d\omega + \int \rho_t\,\nabla(\omega \cdot \Omega^\varepsilon_{\rho_t}) \cdot \nabla\Big(\frac{|\nabla\varphi_t|^2}{2}\Big)\,d\omega. \end{aligned}$$

Similarly we have

$$\begin{aligned} I_{22} &= \int \rho_t\,\nabla\Big(\frac{|\nabla\varphi_t|^2}{2}\Big) \cdot \Omega^\varepsilon_{\rho_t}\,d\omega \\ &= \int \rho_t\,\nabla(\omega \cdot \Omega^\varepsilon_{\rho_t}) \cdot \nabla\Big(\frac{|\nabla\varphi_t|^2}{2}\Big)\,d\omega, \end{aligned}$$

thus

$$I_{21} + I_{22} = -\int \rho_t\,\nabla\varphi_t\,D^2(\omega \cdot \Omega^\varepsilon_{\rho_t})\,\nabla\varphi_t\,d\omega.$$

Concerning I_{23}, since

$$
\begin{aligned}
\partial_t \Omega^\varepsilon_{\rho_t} &= \frac{\partial_t J_{\rho_t}}{\sqrt{|J_{\rho_t}|^2+\varepsilon}} - \frac{\Omega^\varepsilon_{\rho_t}}{\sqrt{|J_{\rho_t}|^2+\varepsilon}}\,\Omega^\varepsilon_{\rho_t}\cdot\partial_t J_{\rho_t} \\
&= \frac{1}{\sqrt{|J_{\rho_t}|^2+\varepsilon}}\Big(-\int \omega\,\nabla\cdot(\rho_t\nabla\varphi_t)\,d\omega + \Omega^\varepsilon_{\rho_t}\int \Omega^\varepsilon_{\rho_t}\cdot\omega\,\nabla\cdot(\rho_t\nabla\varphi_t)\,d\omega\Big) \\
&= \frac{1}{\sqrt{|J_{\rho_t}|^2+\varepsilon}}\Big(\int \rho_t\,\nabla\varphi_t\,d\omega - \Omega^\varepsilon_{\rho_t}\int \nabla(\Omega^\varepsilon_{\rho_t}\cdot\omega)\cdot\nabla\varphi_t\,\rho_t\,d\omega\Big) \\
&= \frac{1}{\sqrt{|J_{\rho_t}|^2+\varepsilon}}\Big(\int \rho_t\,\nabla\varphi_t\,d\omega - \Omega^\varepsilon_{\rho_t}\int \Omega^\varepsilon_{\rho_t}\cdot\nabla\varphi_t\,\rho_t\,d\omega\Big),
\end{aligned}
$$

we have

$$
\begin{aligned}
I_{23} &= -\frac{1}{\sqrt{|J_{\rho_t}|^2+\varepsilon}}\int \rho_t\,\nabla\varphi\cdot\Big(\int \rho_t\,\nabla\varphi_t\,d\omega - \Omega^\varepsilon_{\rho_t}\int \Omega^\varepsilon_{\rho_t}\cdot\nabla\varphi_t\,\rho_t\,d\omega\Big)\,d\omega \\
&= -\frac{1}{\sqrt{|J_{\rho_t}|^2+\varepsilon}}\Big|\int \nabla\varphi_t\,\rho_t\,d\omega\Big|^2 + \frac{1}{\sqrt{|J_{\rho_t}|^2+\varepsilon}}\Big(\int \Omega^\varepsilon_{\rho_t}\cdot\nabla\varphi_t\,\rho_t\,d\omega\Big)^2.
\end{aligned}
$$

Recalling that $\rho_0=\rho$ and $\varphi_0=\varphi$, this completes the proof of (3.2). $\square$

3.2 Minimizing movements for the free energy

To prove existence of solutions to the regularized problem, we use the time-discrete scheme by Jordan, Kindeleherer and Otto [66] (see also [43]). Hence, in all this section, $\varepsilon>0$ is fixed and, to simplify the notation, we shall not explicitly show the dependence on it.

Given a time step $\tau>0$, for a given initial data $\rho_0\in\mathcal{P}(\mathbb{S}^{d-1})$ we set

$$\rho_0^\tau = \rho_0,$$

and then recursively define ρ_n^τ as a minimizer of

$$\sigma \mapsto \frac{W_2^2(\sigma,\rho_{n-1}^\tau)}{2\tau} + \mathcal{E}^\varepsilon(\rho). \tag{3.8}$$

The existence of a minimizer to (3.8) is guaranteed as follows.

Lemma 3.2. *For a given $\tau > 0$ and $\rho \in \mathcal{P}(\mathbb{S}^{d-1})$, there exists a minimum $\rho_\tau \in \mathcal{P}(\mathbb{S}^{d-1})$ of*

$$\sigma \to \frac{W_2^2(\sigma,\rho)}{2\tau} + \mathcal{E}^\varepsilon(\sigma).$$

Furthermore, the optimal transport map T sending $\rho_\tau \, d\omega$ onto $\rho \, d\omega$ is given by

$$T(\omega) = \exp_\omega \left[\tau \nabla_\omega \big(\log \rho_\tau - \omega \cdot \Omega^\varepsilon_{\rho_\tau} \big)\right]. \tag{3.9}$$

Proof. First of all, the existence of a minimum $\mu_\tau = \rho_\tau \, d\omega$ follows from the fact that $\mathcal{E}^\varepsilon$ is lower semicontinous and bounded from below thanks to

$$\begin{aligned} \mathcal{E}^\varepsilon(\rho_t) &\geq \min_{x>0} x \log x \int_{\mathbb{S}^{d-1}} d\omega - \sqrt{\Big| \int_{\mathbb{S}^{d-1}} \rho_t \, d\omega \Big|^2 + \varepsilon} \\ &\geq |\mathbb{S}^{d-1}| \, e^{-1} \log e^{-1} - \sqrt{1+\varepsilon}. \end{aligned} \tag{3.10}$$

To show (3.9), let φ be a $d^2/2$-convex function such that $\exp_\omega(\nabla_\omega \varphi)$ is the optimal map sending $\rho_\tau \, d\omega$ onto $\rho \, d\omega$. For any smooth vector field X on $\mathbb{S}^{d-1}$, we set

$$\mu_t := \exp(tX)_\# \rho_\tau \, d\omega.$$

Using the minimality of ρ_τ, we get

$$\mathcal{E}^\varepsilon(\mu_t) - \mathcal{E}^\varepsilon(\rho_\tau) + \frac{W_2^2(\mu_t,\rho) - W_2^2(\rho_\tau,\rho)}{2\tau} \geq 0.$$

Then we use Proposition 2.1 and (3.3) to obtain

$$\begin{aligned} &\int_{\mathbb{S}^{d-1}} \nabla_\omega \big(\log \rho_\tau - \omega \cdot \Omega^\varepsilon_{\rho_\tau} \big) \cdot X(\omega) \, \rho \, d\omega - \frac{1}{\tau} \int_{\mathbb{S}^{d-1}} \nabla_\omega \varphi \cdot X(\omega) \, \rho \, d\omega \\ &\quad \geq \limsup_{t \to 0} \left(\mathcal{E}^\varepsilon(\mu_t) - \mathcal{E}^\varepsilon(\rho_\tau) + \frac{W_2^2(\mu_t,\rho) - W_2^2(\rho_\tau,\rho)}{2\tau} \right) \geq 0. \end{aligned}$$

Exchanging X with $-X$, this yields

$$\int_{\mathbb{S}^{d-1}} \tau \nabla_\omega \big(\log \rho_\tau - \omega \cdot \Omega^\varepsilon_{\rho_\tau} \big) \cdot X(\omega)\, \rho \, d\omega = \int_{\mathbb{S}^{d-1}} \nabla_\omega \varphi \cdot X(\omega)\, \rho \, d\omega,$$

and since X is arbitrary we get

$$\nabla_\omega \varphi = \tau \nabla_\omega \big(\log \rho_\tau - \omega \cdot \Omega^\varepsilon_{\rho_\tau} \big),$$

which proves (3.9). $\square$

3.3 Existence of the regularized equation (3.1)

Using the sequence of minimizers defined in the previous section, we define the discrete solution $t \mapsto \rho^\tau(t)$ by

$$\rho^\tau(t) := \rho^\tau_n, \qquad \text{for } t \in [n\tau, (n+1)\tau).$$

We show now the existence of weak solutions to (3.1) as a limit of the discrete solutions ρ^τ as $\tau \to 0$.

Proposition 3.1. *Assume $\rho_0 \in \mathcal{P}(\mathbb{S}^{d-1})$ with $\int_{\mathbb{S}^{d-1}} \rho_0 \log \rho_0 \, d\omega < \infty$. Then, for any sequence $\tau_k \downarrow 0$, up to a subsequence $\rho^{\tau_k}(t)$ converges to some limit $\rho(t)$ locally uniformly in time. The limit $t \mapsto \rho(t)$ belongs to $L^2_{loc}([0,\infty), W^{1,1}(\mathbb{S}^{d-1}))$ and is a weak solution of* (3.1)*.*

Proof. Throughout the proof, we will use the following inequality for the sequence of minimizers ρ^τ_n of (3.8),

$$\frac{1}{2} \sum_{i=n}^{m-1} \frac{W_2^2(\rho^\tau_{i+1}, \rho^\tau_i)}{\tau} + \frac{\tau}{2} \sum_{i=n}^{m-1} |\nabla \mathcal{E}^\varepsilon(\rho^\tau_i)|^2 \leq \mathcal{E}^\varepsilon(\rho^\tau_m) - \mathcal{E}^\varepsilon(\rho^\tau_n), \qquad \text{for any } n < m, \tag{3.11}$$

referring to [6, Lemma 3.2.2] for its proof.

Since $\mathcal{E}^\varepsilon(\rho_m^\tau) \leq \mathcal{E}^\varepsilon(\rho_0)$ for all m and $\mathcal{E}^\varepsilon(\rho_n^\tau)$ bounded from below due to (3.10), we have

$$\mathcal{E}^\varepsilon(\rho_m^\tau) - \mathcal{E}^\varepsilon(\rho_n^\tau) \leq \mathcal{E}^\varepsilon(\rho_0) + \sqrt{1+\varepsilon}. \tag{3.12}$$

Let $\{\tau_k\}_{k\in\mathbb{N}}$ be a sequence converging to 0. Then, for any $n < m$,

$$\frac{1}{2}\sum_{i=n}^{m-1} \frac{W_2^2(\rho_{i+1}^{\tau_k}, \rho_i^{\tau_k})}{\tau_k} \leq \mathcal{E}^\varepsilon(\rho_0) + \sqrt{2} \tag{3.13}$$

(recall that $\varepsilon \leq 1$).

Notice that

$$\mathcal{E}^\varepsilon(\rho_0) \leq \int_{\mathbb{S}^{d-1}} \rho_0 \log \rho_0 \, d\omega < \infty. \tag{3.14}$$

Also, it follows by Jensen's inequality that

$$\begin{aligned}\frac{W_2^2(\rho_m^\tau, \rho_n^\tau)}{(m-n)^2} \leq \Big(\frac{\sum_{i=n}^{m-1} W_2(\rho_{i+1}^\tau, \rho_i^\tau)}{m-n}\Big)^2 \\ \leq \frac{\sum_{i=n}^{m-1} W_2^2(\rho_{i+1}^\tau, \rho_i^\tau)}{m-n} \leq 2\tau(\mathcal{E}^\varepsilon(\rho_0) + \sqrt{2}).\end{aligned}$$

Hence, setting $n = [\frac{s}{\tau}]$ and $m = [\frac{t}{\tau}]$ for any $0 \leq s < t$, we have

$$W_2(\rho^{\tau_k}(t), \rho^{\tau_k}(s)) \leq \sqrt{2(\mathcal{E}^\varepsilon(\rho_0) + \sqrt{2})\,[t - s + \tau_k]}. \tag{3.15}$$

This equicontinuity estimate combined the compactness of $(\mathcal{P}(\mathbb{S}^{d-1}), W_2)$ implies that, up to a subsequence,

$$\rho^{\tau_k}(t) \text{ converges to some limit } \rho(t) \text{ in } (\mathcal{P}(\mathbb{S}^{d-1}), W_2) \text{ locally uniformly in } t \geq 0. \tag{3.16}$$

We now show that $t \mapsto \rho(t)$ is a weak solution of (3.1). For $n \in \mathbb{N}$, by (3.8) and Lemma 3.2, we have

$$\Big(\exp_\omega \big(\tau_k \nabla_\omega \big(\log \rho_{n+1}^{\tau_k} - \omega \cdot \Omega^{\varepsilon}_{\rho_{n+1}^{\tau_k}} \big) \big) \Big)_{\#} \rho_{n+1}^{\tau_k} \, d\omega = \rho_n^{\tau_k} \, d\omega.$$

Thus, for any $\varphi \in C^\infty(\mathbb{S}^{d-1})$,

$$\begin{aligned} &\int_{\mathbb{S}^{d-1}} \varphi(\omega) \, (\rho_{n+1}^{\tau_k} - \rho_n^{\tau_k}) \, d\omega \\ &\qquad = \int_{\mathbb{S}^{d-1}} \Big(\varphi(\omega) - \varphi\big(\exp_\omega(\tau_k \nabla_\omega (\log \rho_{n+1}^{\tau_k} - \omega \cdot \Omega^{\varepsilon}_{\rho_{n+1}^{\tau_k}})) \big) \Big) \, \rho_{n+1}^{\tau_k} \, d\omega. \end{aligned}$$

Using, for each $\omega \in \mathbb{S}^{d-1}$, the Taylor formula along the geodesic $s \mapsto \exp_\omega(s \tau_k \nabla_\omega (\log \rho_{n+1}^{\tau_k} - \omega \cdot \Omega^{\varepsilon}_{\rho_{n+1}^{\tau_k}}))$, we have

$$\begin{aligned} &\int_{\mathbb{S}^{d-1}} \varphi(\omega) \, (\rho_{n+1}^{\tau_k} - \rho_n^{\tau}) \, d\omega \\ &= - \int_{\mathbb{S}^{d-1}} \tau_k \, \nabla_\omega \varphi(\omega) \cdot \nabla_\omega \big[\log \rho_{n+1}^{\tau_k} - \omega \cdot \Omega^{\varepsilon}_{\rho_{n+1}^{\tau_k}} \big] \, \rho_{n+1}^{\tau_k} \, d\omega + R(n, \tau_k) \qquad (3.17) \\ &= \int_{\mathbb{S}^{d-1}} \tau_k \big[\Delta_\omega \varphi + \nabla_\omega \varphi \cdot \nabla_\omega (\omega \cdot \Omega^{\varepsilon}_{\rho_{n+1}^{\tau_k}}) \big] \rho_{n+1}^{\tau_k} \, d\omega + R(n, \tau_k), \end{aligned}$$

where the remainder term $R(n, \tau_k)$ can be estimated by

$$\begin{aligned} &R(n,\tau_k) \leq \|D^2_\omega \varphi\|_{L^\infty(\mathbb{S}^{d-1})} \int_{S^{d-1}} d^2 \Big(\omega, \exp_\omega(\tau_k \nabla_\omega \big[\log \rho_{n+1}^{\tau_k} - \omega \cdot \Omega_{\rho_{n+1}^{\tau_k}} \big] \Big) \, \rho_{n+1}^{\tau_k} \, d\omega \\ &= \|D^2_\omega \varphi\|_{L^\infty(\mathbb{S}^{d-1})} \, W_2^2(\rho_{n+1}^{\tau_k}, \rho_n^{\tau_k}), \end{aligned} \qquad (3.18)$$

For any $0 \le t < s$, we sum up (3.17) from $l := [\frac{t}{\tau_k}]$ to $m := [\frac{s}{\tau_k}]$ to get

$$\begin{aligned}
&\int_{\mathbb{S}^{d-1}} \varphi\left(\rho^{\tau_k}(s) - \rho^{\tau_k}(t)\right) d\omega \\
&\qquad \le \sum_{n=l}^{m} \int_{\mathbb{S}^{d-1}} \tau_k \left[\Delta_\omega \varphi + \nabla_\omega \varphi \cdot \nabla_\omega (\omega \cdot \Omega^\varepsilon_{\rho^{\tau_k}_{n+1}})\right] \rho^{\tau_k}_{n+1}\, d\omega \\
&\qquad\quad + \sum_{n=l}^{m} R(n, \tau_k) \\
&\qquad = \int_{(l+1)\tau_k}^{(m+2)\tau_k} \int_{\mathbb{S}^{d-1}} \left[\Delta_\omega \varphi + \nabla_\omega \varphi \cdot \nabla_\omega (\omega \cdot \Omega^\varepsilon_{\rho^{\tau_k}(r)})\right] \rho^{\tau_k}(r)\, d\omega\, dr \\
&\qquad\quad + \sum_{n=l}^{m} R(n, \tau_k).
\end{aligned}$$

Letting $\tau_k \to 0$, (3.16) implies

$$\int_{\mathbb{S}^{d-1}} \varphi\left(\rho^{\tau_k}(s) - \rho^{\tau_k}(t)\right) d\omega \to \int_{\mathbb{S}^{d-1}} \varphi\left(\rho(s) - \rho(t)\right) d\omega.$$

Since $J_{\rho^{\tau_k}} \to J_\rho$, we have

$$\begin{aligned}
|\Omega^\varepsilon_{\rho^{\tau_k}} - \Omega^\varepsilon_\rho| &\le \frac{\left| J_{\rho^{\tau_k}}\left(\sqrt{|J_\rho|^2+\varepsilon} - \sqrt{|J_{\rho^{\tau_k}}|^2+\varepsilon}\right) + \sqrt{|J_{\rho^{\tau_k}}|^2+\varepsilon}\,(J_{\rho^{\tau_k}} - J_\rho)\right|}{\sqrt{|J_{\rho^{\tau_k}}|^2+\varepsilon}\sqrt{|J_\rho|^2+\varepsilon}} \\
&\le \frac{1}{\varepsilon}\Big(|J_{\rho^{\tau_k}} - J_\rho| + \Big|\sqrt{|J_{\rho^{\tau_k}}|^2+\varepsilon} - \sqrt{|J_\rho|^2+\varepsilon}\Big|\Big) \\
&\to 0,
\end{aligned}$$

which implies that, for all r,

$$\begin{aligned}
&\int_{\mathbb{S}^{d-1}} \left[\Delta_\omega \varphi + \nabla_\omega \varphi \cdot \nabla_\omega (\omega \cdot \Omega^\varepsilon_{\rho^{\tau_k}(r)})\right] \rho^{\tau_k}(r)\, d\omega \\
&\qquad \to \int_{\mathbb{S}^{d-1}} \left[\Delta_\omega \varphi + \nabla_\omega \varphi \cdot \nabla_\omega (\omega \cdot \Omega^\varepsilon_{\rho(r)})\right] \rho(r)\, d\omega.
\end{aligned}$$

Moreover since $\int_{\mathbb{S}^{d-1}} \left[\Delta_\omega \varphi + \nabla_\omega \varphi \cdot \nabla_\omega (\omega \cdot \Omega^\varepsilon_{\rho^{\tau_k}(r)})\right] \rho^{\tau_k}(r)\, d\omega$ is uniformly bounded, the dominated convergence theorem yields

$$\int_{(l+1)\tau_k}^{(m+2)\tau_k} \int_{\mathbb{S}^{d-1}} \left[\Delta_\omega \varphi + \nabla_\omega \varphi \cdot \nabla_\omega (\omega \cdot \Omega^\varepsilon_{\rho^{\tau_k}(r)})\right] \rho^{\tau_k}(r)\, d\omega\, dr$$
$$\to \int_t^s \int_{\mathbb{S}^{d-1}} \left[\Delta_\omega \varphi + \nabla_\omega \varphi \cdot \nabla_\omega (\omega \cdot \Omega^\varepsilon_{\rho(r)})\right] \rho(r)\, d\omega\, dr.$$

On the other hand, since (3.18) and (3.13) give

$$\sum_{n=l}^{m} R(n, \tau_k) \le C \sum_{n=l}^{m} W_2^2(\rho^{\tau_k}_{n+1}, \rho^{\tau_k}_n)$$
$$\le C(\mathcal{E}^\varepsilon(\rho_0) + \sqrt{2})\tau_k \to 0,$$

we have shown that $0 \le t < s$,

$$\int_{\mathbb{S}^{d-1}} \varphi\,(\rho(s) - \rho(t))\, d\omega = \int_t^s \int_{\mathbb{S}^{d-1}} \left[\Delta_\omega \varphi + \nabla_\omega \varphi \cdot \nabla_\omega (\omega \cdot \Omega^\varepsilon_{\rho(r)})\right] \rho(r)\, d\omega\, dr,$$

which provides the weak formulation of (3.1). Moreover, thanks to (3.15) and (3.14),

$$W_2(\rho(t), \rho(s)) \le \sqrt{2 \int_{\mathbb{S}^{d-1}} \rho_0 \log \rho_0\, d\omega}\, \sqrt{t - s}, \tag{3.19}$$

hence $t \mapsto \rho(t)$ is weakly continuous and ρ is a weak solution to (3.1).

It remains to show that $\rho \in L^2_{loc}([0, \infty), W^{1,1}(\mathbb{S}^{d-1}))$. Using again (3.11) and (3.12) we see that, for any $0 \le t < s$,

$$\int_t^s |\nabla \mathcal{E}^\varepsilon(\rho^{\tau_k}(t))|^2 dt \le 2(\mathcal{E}^\varepsilon(\rho_0) + \sqrt{2}),$$

which together with (3.4) yields

$$
\begin{aligned}
2(\mathcal{E}^\varepsilon(\rho_0)+\sqrt{2}) &\geq \int_t^s \int_{\mathbb{S}^{d-1}} |\nabla_\omega(\log \rho^{\tau_k} - \omega\cdot\Omega^\varepsilon_{\rho^{\tau_k}})|^2 \rho^{\tau_k}\, d\omega\, dt \\
&\geq \int_t^s \int_{\mathbb{S}^{d-1}} |\nabla_\omega \log \rho^{\tau_k}|^2 \rho^{\tau_k}\, d\omega\, dt \\
&\quad - 2\int_t^s \int_{\mathbb{S}^{d-1}} \nabla_\omega \log \rho^{\tau_k}\cdot \nabla_\omega(\omega\cdot\Omega^\varepsilon_{\rho^{\tau_k}})\, \rho^{\tau_k}\, d\omega\, dt \qquad (3.20)\\
&\geq \frac{1}{2}\int_t^s \int_{\mathbb{S}^{d-1}} |\nabla_\omega \log \rho^{\tau_k}|^2 \rho^{\tau_k}\, d\omega\, dt \\
&\quad - 2\int_t^s \int_{\mathbb{S}^{d-1}} |\nabla_\omega(\omega\cdot\Omega^\varepsilon_{\rho^{\tau_k}})|^2 \rho^{\tau_k}\, d\omega\, dt.
\end{aligned}
$$

Since $|\nabla_\omega(\omega\cdot\Omega^\varepsilon)| = |\mathbb{P}_{\omega^\perp}\Omega| \leq 1$, we have

$$
\begin{aligned}
\int_t^s \int_{\mathbb{S}^{d-1}} |\nabla_\omega \sqrt{\rho^{\tau_k}}|^2 d\omega\, dt &= \frac{1}{2}\int_t^s \int_{\mathbb{S}^{d-1}} |\nabla_\omega \log \rho^{\tau_k}|^2 \rho^{\tau_k}\, d\omega\, dt \\
&\leq 2(\mathcal{E}^\varepsilon(\rho_0)+\sqrt{2}) + 2(s-t),
\end{aligned}
$$

which implies that $\sqrt{\rho^{\tau_k}}$ is uniformly bounded in $L^2_{loc}([0,+\infty), H^1(\mathbb{S}^{d-1}))$. Therefore, letting $\tau_k \to 0$, we get

$$
\sqrt{\rho} \in L^2_{loc}([0,+\infty), H^1(\mathbb{S}^{d-1})), \tag{3.21}
$$

that combined with Hölder inequality implies that $\rho \in L^2_{loc}([0,+\infty), W^{1,1}(\mathbb{S}^{d-1}))$.

□

3.4 Uniqueness

The following results provide the stability estimates for weak solutions to (3.1), thus their uniqueness. We shall revisit the arguments of the proof to show the stability and uniqueness of weak solutions to (1.1) in Section 5.

Proposition 3.2. *(Uniqueness and stability). Assume $\rho_0, \bar\rho_0 \in \mathcal{P}(\mathbb{S}^{d-1})$ satisfy* (2.7)*. Let $\rho^\varepsilon, \bar\rho^\varepsilon$ be solutions of* (3.1) *with corresponding initial datas $\rho_0, \bar\rho_0$. Then for all $t > 0$,*

$$W_2(\rho^\varepsilon(t), \bar\rho^\varepsilon(t)) \leq e^{\lambda t} W_2(\rho_0, \bar\rho_0), \tag{3.22}$$

where $\lambda := (1 + \varepsilon^{-1/2}) - (d-2)$.

Proof. For a fixed time $t > 0$, let φ_0 be a $d^2/2$-convex function such that $\exp_\omega(\nabla_\omega \varphi_0)$ is the optimal map sending $\rho^\varepsilon(t)\, d\omega$ onto $\bar\rho^\varepsilon(t)\, d\omega$, and consider the curve $[0,1] \ni r \mapsto \alpha_r\, d\omega$ of absolutely continuous measures defined by

$$\alpha_r\, d\omega = \exp_\omega(r\nabla_\omega \varphi_0)_\# \rho^\varepsilon(t)\, d\omega$$

(the absolute continuity of α_r follows, for instance, from [40, Section 5]). Then the curve $r \mapsto \alpha_r\, d\omega$ is the unique geodesic in $(\mathcal{P}(\mathbb{S}^{d-1}), W_2)$ connecting $\alpha_0 = \rho^\varepsilon(t)$ to $\alpha_1 = \bar\rho^\varepsilon(t)$ (see for example [5, Corollary 3.22]).
For each $r \in [0,1]$, let φ_r be a $d^2/2$-convex function such that $\exp_\omega(\nabla_\omega \varphi_r)$ is the optimal map sending $\alpha_r\, d\omega$ onto $\bar\rho^\varepsilon(t)\, d\omega$. Similarly, the curve $s \mapsto \alpha_{r,s}\, d\omega$ defined by

$$\alpha_{r,s}\, d\omega = \exp_\omega(s\nabla\varphi_r)_\# \alpha_r\, d\omega, \tag{3.23}$$

and it is the unique geodesic in $(\mathcal{P}(\mathbb{S}^{d-1}), W_2)$ connecting $\alpha_{r,0} = \alpha_r$ to $\alpha_{r,1} = \bar\rho^\varepsilon(t)$. Notice that it follows from the uniqueness of the geodesics that, for all $r, s \in [0,1]$,

$$\alpha_{r+(1-r)s} = \alpha_{r,s}.$$

Now, applying (3.2) in Lemma 3.1 to (3.23), we estimate the second derivative of $\mathcal{E}^\varepsilon$ by Wasserstein distance as

$$\begin{aligned}
\frac{d^2}{dh^2}\bigg|_{h=r} \mathcal{E}^\varepsilon(\alpha_h) &= \frac{d^2}{dh^2}\bigg|_{h=0} \mathcal{E}^\varepsilon(\alpha_{r,\frac{h}{1-r}}) \\
&= \frac{1}{(1-r)^2}\frac{d^2}{ds^2}\bigg|_{s=0} \mathcal{E}^\varepsilon(\alpha_{r,s}) \\
&\geq -\frac{\lambda}{(1-r)^2}\int_{\mathbb{S}^{d-1}} |\nabla\varphi_r|^2 \alpha_r \, d\omega \\
&= -\lambda \, \frac{W_2^2(\alpha_r, \bar{\rho}^\varepsilon(t))}{(1-r)^2} \\
&= -\lambda \, W_2^2(\rho^\varepsilon(t), \bar{\rho}^\varepsilon(t)),
\end{aligned} \tag{3.24}$$

where $\lambda := (1+\varepsilon^{-1/2}) - (d-2)$ (see (3.5)).

Since, by Taylor formula along the geodesic $r \mapsto \alpha_r \, d\omega$,

$$\mathcal{E}^\varepsilon(\alpha_1) = \mathcal{E}^\varepsilon(\alpha_0) + \frac{d}{dr}\bigg|_{r=0} \mathcal{E}^\varepsilon(\alpha_r) + \int_0^1 (1-r)\frac{d^2}{dr^2}\mathcal{E}^\varepsilon(\alpha_r)\, dr,$$

we use (3.3) and (3.24) to have

$$\begin{aligned}
&\mathcal{E}^\varepsilon(\bar{\rho}^\varepsilon(t)) \geq \mathcal{E}^\varepsilon(\rho^\varepsilon(t)) \\
&\qquad + \int_{\mathbb{S}^{d-1}} \nabla\varphi_0 \cdot \nabla(\log\rho^\varepsilon(t) - \omega\cdot\Omega_{\rho^\varepsilon(t)})\, \rho^\varepsilon(t)\, d\omega - \frac{\lambda}{2}\, W_2^2(\rho^\varepsilon(t), \bar{\rho}^\varepsilon(t)).
\end{aligned}$$

Similarly, applying the above arguments to the $d^2/2$-convex function $\overline{\varphi}_0$ satisfying

$$\rho^\varepsilon(t)\, d\omega = \exp_\omega(\nabla\overline{\varphi}_0)_\# \bar{\rho}^\varepsilon(t)\, d\omega,$$

we have

$$\begin{aligned}
&\mathcal{E}^\varepsilon(\rho^\varepsilon(t)) \geq \mathcal{E}^\varepsilon(\bar{\rho}^\varepsilon(t)) \\
&\qquad + \int_{\mathbb{S}^{d-1}} \nabla\bar{\varphi}_0 \cdot \nabla(\log\bar{\rho}^\varepsilon(t) - \omega\cdot\Omega_{\bar{\rho}^\varepsilon(t)})\, \bar{\rho}^\varepsilon(t)\, d\omega - \frac{\lambda}{2}\, W_2^2(\rho^\varepsilon(t), \bar{\rho}^\varepsilon(t)),
\end{aligned}$$

therefore

$$
\begin{aligned}
&\int_{\mathbb{S}^{d-1}} \nabla\varphi_0 \cdot \nabla(\log\rho^\varepsilon(t) - \omega\cdot\Omega_{\rho^\varepsilon(t)})\,\rho^\varepsilon(t)\,d\omega \\
&\quad + \int_{\mathbb{S}^{d-1}} \nabla\bar\varphi_0 \cdot \nabla(\log\bar\rho^\varepsilon(t) - \omega\cdot\Omega_{\bar\rho^\varepsilon(t)})\,\bar\rho^\varepsilon(t)\,d\omega \le \lambda\, W_2^2(\rho^\varepsilon(t),\bar\rho^\varepsilon(t)).
\end{aligned} \tag{3.25}
$$

We now claim that

$$
\begin{aligned}
\frac{d}{dt}W_2^2(\rho^\varepsilon(t),\bar\rho^\varepsilon(t)) = &\int_{\mathbb{S}^{d-1}} \nabla\varphi_0 \cdot \nabla(\log\rho^\varepsilon(t) - \omega\cdot\Omega_{\rho^\varepsilon(t)})\,\rho^\varepsilon(t)\,d\omega \\
&+ \int_{\mathbb{S}^{d-1}} \nabla\bar\varphi_0 \cdot \nabla(\log\bar\rho^\varepsilon(t) - \omega\cdot\Omega_{\bar\rho^\varepsilon(t)})\,\bar\rho^\varepsilon(t)\,d\omega.
\end{aligned} \tag{3.26}
$$

Indeed, ρ^ε and $\bar\rho^\varepsilon$ solve the continuity equation

$$
\partial_t\rho + \nabla_\omega\cdot(v[\rho]\rho) = 0,
$$

where $v[\rho] := \nabla_\omega(\omega\cdot\Omega_\rho^\varepsilon - \log\rho)$ is a locally Lipschitz vector field. Moreover it follows from (3.21) that, for all $t < s$,

$$
\int_t^s \int_{\mathbb{S}^{d-1}} |v[\rho]|\,\rho\,d\omega \le C\,(s-t)\Big(1 + \|\nabla\sqrt{\rho}\|_{L^2(\mathbb{S}^{d-1})}\Big) < \infty.
$$

Hence the hypotheses of [94, Theorem 23.9] are satisfied implying (3.26), and combining it with (3.25) yields

$$
\frac{d}{dt}W_2^2(\rho^\varepsilon(t),\bar\rho^\varepsilon(t) \le \lambda\, W_2^2(\rho^\varepsilon(t),\bar\rho^\varepsilon(t)),
$$

which completes the proof. □

3.5 Properties of solutions to (3.1)

In the next lemma, we show that the momentum does not vanish for any finite time $t \in (0,\infty)$, with an estimate independent of ε.

Lemma 3.3. *Let ρ^ε be a solution of* (3.1)*. Then, for all $t > 0$,*

$$|J_{\rho^\varepsilon(t)}|^2 \geq |J_{\rho_0}|^2 e^{-2(d-1)t}.$$

Proof. It follows from (3.1) that

$$\begin{aligned}
\frac{d}{dt}\frac{1}{2}|J_{\rho^\varepsilon}|^2 &= J_{\rho^\varepsilon} \cdot \partial_t J_{\rho^\varepsilon} \\
&= J_{\rho^\varepsilon} \cdot \left(\int \omega\, \Delta\rho^\varepsilon\, d\omega - \int \omega\, \nabla \cdot (\rho^\varepsilon \nabla(\omega \cdot \Omega_{\rho^\varepsilon}))\, d\omega \right) \\
&= J_{\rho^\varepsilon} \cdot \int \omega\, \Delta\rho^\varepsilon\, d\omega - \int J_{\rho^\varepsilon} \cdot \omega\, \nabla \cdot (\rho^\varepsilon \nabla(\omega \cdot \Omega_{\rho^\varepsilon}))\, d\omega \\
&=: I_1 + I_2.
\end{aligned}$$

We use (2.4) to get

$$\begin{aligned}
I_1 &= -J_{\rho^\varepsilon} \cdot \int \nabla\rho^\varepsilon\, d\omega \\
&= -(d-1) J_{\rho^\varepsilon} \cdot \int \omega\, \rho^\varepsilon\, d\omega \\
&= -(d-1)|J_{\rho^\varepsilon}|^2.
\end{aligned}$$

Also, using (2.3), we have

$$\begin{aligned}
I_2 &= \int \nabla(J_{\rho^\varepsilon} \cdot \omega) \cdot \nabla(\omega \cdot \Omega_{\rho^\varepsilon})\, \rho^\varepsilon d\omega \\
&= \int \frac{\rho^\varepsilon}{\sqrt{|J_{\rho^\varepsilon}|^2 + \varepsilon}}\, |\nabla(\omega \cdot J_{\rho^\varepsilon})|^2\, d\omega \\
&\geq 0.
\end{aligned}$$

Thus

$$\frac{d}{dt}|J_{\rho^\varepsilon}|^2 \leq -2(d-1)|J_{\rho^\varepsilon}|^2,$$

which completes the proof. □

3.6 Proof of the existence in Theorem 2.1

Let $\{\varepsilon_k\}_{k\in\mathbb{N}}$ be a sequence converging to 0. As a consequence of (3.19) it follows that the sequence $\{\rho^{\varepsilon_k}\}_{k\in\mathbb{N}}$ is equicontinous, so the compactness of $(\mathcal{P}(\mathbb{S}^{d-1}), W_2)$ imply that up to a subsequence, $\rho^{\varepsilon_k}(t)$ converges to some limit $\rho(t)$ in $(\mathcal{P}(\mathbb{S}^{d-1}), W_2)$ uniformly in $t \geq 0$. Then since $J_{\rho^{\varepsilon_k}} \to J_\rho$, it follows from Lemma 3.3 that for all $t > 0$,

$$|J_{\rho(t)}|^2 \geq |J_{\rho_0}|^2 e^{-2(d-1)t}. \tag{3.27}$$

Therefore by the same arguments as the proof of Proposition 3.1, the limit ρ is a weak solution to (1.1). Moreover since a straightforward computation yields

$$\frac{d}{dt}\mathcal{E}^0(\rho) = -\int_{\mathbb{S}^{d-1}} |\nabla_\omega(\log\rho - \omega\cdot\Omega_\rho)|^2 \rho\, d\omega,$$

the analogue of (3.20) with $\varepsilon = 0$ combined with (3.14) provide

$$\rho \in L^2_{loc}([0,+\infty), W^{1,1}(\mathbb{S}^{d-1})).$$

$\square$

4 Convergence towards equilibrium

In this section, we prove Theorem 2.2. We start with the following estimates on the difference between ρ^ε and $M_{\Omega^\varepsilon_{\rho^\varepsilon}}$.

Lemma 4.1. *Let C_M be as in (1.3), and let ρ^ε be a solution of (3.1) starting from ρ_0. Then, for all $t > 0$,*

$$\|\rho^\varepsilon(t) - M_{\Omega^\varepsilon_{\rho^\varepsilon(t)}}\|_{L^1(\mathbb{S}^{d-1})} \leq e^{-C_1 t}\Big(\int_{\mathbb{S}^{d-1}} \rho_0 \log\rho_0\, d\omega + 1 - \log C_M\Big) + \sqrt{\varepsilon}.$$

where

$$C_1 := \frac{2(d-2)}{e^2}.$$

Proof. First of all, for each measure ρ^ε, we denote its relative entropy with respect to the probability measure $M_{\Omega^\varepsilon_{\rho^\varepsilon}}(\omega)\,d\omega = C_M e^{\omega\cdot\Omega^\varepsilon_{\rho^\varepsilon}}\,d\omega$ by

$$H(\rho^\varepsilon \mid M_{\Omega^\varepsilon_{\rho^\varepsilon}}) = \int_{\mathbb{S}^{d-1}} \rho^\varepsilon \log\Big(\frac{\rho^\varepsilon}{M_{\Omega^\varepsilon_{\rho^\varepsilon}}}\Big)\,d\omega,$$

which can also be rewritten as

$$H(\rho^\varepsilon \mid M_{\Omega^\varepsilon_{\rho^\varepsilon}}) = \int_{\mathbb{S}^{d-1}} \rho^\varepsilon \log \rho^\varepsilon\,d\omega - \int_{\mathbb{S}^{d-1}} \omega\cdot\Omega^\varepsilon_{\rho^\varepsilon}\,\rho^\varepsilon\,d\omega - \log C_M.$$

Since

$$\begin{aligned}\int_{\mathbb{S}^{d-1}} \omega\cdot\Omega^\varepsilon_{\rho^\varepsilon}\,\rho^\varepsilon\,d\omega &= \frac{J_{\rho^\varepsilon}}{\sqrt{|J_{\rho^\varepsilon}|^2+\varepsilon}}\cdot\int \omega\,\rho^\varepsilon\,d\omega = \frac{|J_{\rho^\varepsilon}|^2}{\sqrt{|J_{\rho^\varepsilon}|^2+\varepsilon}}\\ &= \sqrt{|J_{\rho^\varepsilon}|^2+\varepsilon} - \frac{\varepsilon}{\sqrt{|J_{\rho^\varepsilon}|^2+\varepsilon}},\end{aligned}$$

we have

$$\mathcal{E}^\varepsilon(\rho^\varepsilon) = H(\rho^\varepsilon \mid M_{\Omega^\varepsilon_{\rho^\varepsilon}}) - \frac{\varepsilon}{\sqrt{|J_{\rho^\varepsilon}|^2+\varepsilon}} + \log C_M. \tag{4.1}$$

We now set $\alpha := |\mathbb{S}^{d-1}|^{-1}$, and regard the measure $M_{\Omega^\varepsilon_{\rho^\varepsilon}}\,d\omega$ as a bounded perturbation of the constant probability measure $\alpha\,d\omega$, i.e.,

$$M_{\Omega^\varepsilon_{\rho^\varepsilon}} = e^{\omega\cdot\Omega^\varepsilon_{\rho^\varepsilon} - \log C_M} = e^{\omega\cdot\Omega^\varepsilon_{\rho^\varepsilon} - \log C_M - \log\alpha}\alpha,$$

where

$$\operatorname{osc}(\omega\cdot\Omega^\varepsilon_{\rho^\varepsilon} - \log C_M - \log\alpha) \le 2. \tag{4.2}$$

Since the Ricci curvature tensor of $\mathbb{S}^{d-1}$ is $(d-2)I_d$ and $d \geq 3$, the logarithmic Sobolev inequality [10] implies

$$H(\rho^\varepsilon \mid \alpha) \leq \frac{1}{2(d-2)} \int \left|\nabla \log \frac{\rho^\varepsilon}{\alpha}\right|^2 \rho^\varepsilon \, d\omega.$$

Thus, since the logarithmic Sobolev inequality is stable under bounded perturbations (see for instance [63, 84]), it follows from (4.2) that

$$H(\rho^\varepsilon \mid M_{\Omega^\varepsilon_{\rho^\varepsilon}}) \leq \frac{e^2}{2(d-2)} \int \left|\nabla \log \frac{\rho^\varepsilon}{M_{\Omega^\varepsilon_{\rho^\varepsilon}}}\right|^2 \rho^\varepsilon \, d\omega. \tag{4.3}$$

Therefore, since (4.1) yields

$$\frac{d}{dt}\mathcal{E}^\varepsilon(\rho^\varepsilon) = \frac{d}{dt} H(\rho^\varepsilon \mid M_{\Omega^\varepsilon_{\rho^\varepsilon}}) - \frac{d}{dt}\frac{\varepsilon}{\sqrt{|J_{\rho^\varepsilon}|^2 + \varepsilon}},$$

and we see

$$\begin{aligned} \frac{d}{dt}\mathcal{E}^\varepsilon(\rho^\varepsilon) &= -\int |\nabla(\log \rho^\varepsilon - \omega \cdot \Omega^\varepsilon_{\rho^\varepsilon})|^2 \rho^\varepsilon \, d\omega \\ &= -\int \left|\nabla \log \frac{\rho^\varepsilon}{M_{\Omega^\varepsilon_{\rho^\varepsilon}}}\right|^2 \rho^\varepsilon \, d\omega, \end{aligned} \tag{4.4}$$

it follows from (4.3) and (4.4) that

$$\frac{d}{dt} H(\rho^\varepsilon \mid M_{\Omega^\varepsilon_{\rho^\varepsilon}}) \leq -\frac{2(d-2)}{e^2} H(\rho^\varepsilon \mid M_{\Omega^\varepsilon_{\rho^\varepsilon}}) + \frac{d}{dt}\frac{\varepsilon}{\sqrt{|J_{\rho^\varepsilon}|^2 + \varepsilon}}.$$

Integrating this inequality, we get

$$\begin{aligned} &H(\rho^\varepsilon(t) \mid M_{\Omega^\varepsilon_{\rho^\varepsilon(t)}}) \leq e^{-C_1 t} H(\rho_0 \mid M_{\Omega^\varepsilon_{\rho_0}}) + e^{-C_1 t}\int_0^t e^{C_1 s}\frac{d}{ds}\frac{\varepsilon}{\sqrt{|J_{\rho^\varepsilon(s)}|^2 + \varepsilon}}\, ds \\ &\quad = e^{-C_1 t} H(\rho_0 \mid M_{\Omega^\varepsilon_{\rho_0}}) + \frac{\varepsilon}{\sqrt{|\ J_{\rho^\varepsilon(t)}\ |^2 + \varepsilon}} - e^{-C_1 t}\frac{\varepsilon}{\sqrt{|\ J_{\rho_0}\ |^2 + \varepsilon}} \\ &\qquad - C_1 e^{-C_1 t}\int_0^t e^{C_1 s}\frac{\varepsilon}{\sqrt{|J_{\rho^\varepsilon(s)}|^2 + \varepsilon}}\, ds \\ &\quad \leq e^{-C_1 t} H(\rho_0 \mid M_{\Omega^\varepsilon_{\rho_0}}) + \sqrt{\varepsilon}. \end{aligned}$$

Hence, thanks to the Csiszar-Kullback-Pinsker inequality (see for example [51, Theorem 1.4]) and the bound

$$H(\rho_0 \mid M_{\Omega^\varepsilon_{\rho_0}}) \leq \int_{\mathbb{S}^{d-1}} \rho_0 \log \rho_0 \, d\omega + 1 - \log C_M,$$

we have the desired inequality. □

The above estimates immediately imply that our weak solutions of (1.1) looks more and more as a Fisher-von Mises distribution as $t \to \infty$.

Proposition 4.1. *Let ρ be a solution of* (1.1)*. Then for all $t > 0$,*

$$\|\rho(t) - M_{\Omega_{\rho(t)}}\|_{L^1(\mathbb{S}^{d-1})} \leq e^{-\frac{2(d-2)}{e^2}t} \Big(\int_{\mathbb{S}^{d-1}} \rho_0 \log \rho_0 \, d\omega + 1 - \log C_M \Big).$$

Proof. The desired inequality follows by taking $\varepsilon \to 0$ in Lemma 4.1 for each $t > 0$. □

The above proposition only tells us that our solution $\rho(t)$ resembles to $M_{\Omega_{\rho(t)}}$ for $t \gg 1$, but it does not say whether the vector $\Omega_{\rho(t)}$ stabilizes to a fixed vector as $t \to \infty$.

To prove this fact, we first use the above result to obtain the uniform positivity of $|J_\rho|$ in time, which improves the estimate (3.27). The following result ensures that if there is a limit J_∞ of $J_{\rho(t)}$ as $t \to \infty$, then J_∞ has to be a nonzero vector.

Lemma 4.2. *Let ρ be a solution of* (1.1) *with initial data $\rho_0 \in \mathcal{P}(\mathbb{S}^{d-1})$ satisfying* (2.7)*. Then there exists a positive constant $C(\rho_0)$ only depending on ρ_0 such that for all $t > 0$,*

$$|J_{\rho(t)}| > C(\rho_0).$$

Proof. Using Proposition 4.1, we have

$$\begin{aligned}\left|J_{\rho(t)} - \int_{\mathbb{S}^{d-1}} \omega \cdot M_{\Omega_{\rho(t)}}\, d\omega\right| &= \left|\int_{\mathbb{S}^{d-1}} \omega \left(\rho(t) - M_{\Omega_{\rho(t)}}\right) d\omega\right| \\ &\le \|\rho(t) - M_{\Omega_{\rho(t)}}\|_{L^1(\mathbb{S}^{d-1})} \\ &\le e^{-\frac{2(d-2)}{e^2}t}\Big(\int_{\mathbb{S}^{d-1}} \rho_0 \log \rho_0\, d\omega + 1 - \log C_M\Big),\end{aligned}$$

which yields

$$\begin{aligned}|J_{\rho(t)}| &\ge \left|\int_{\mathbb{S}^{d-1}} \omega \cdot M_{\Omega_{\rho(t)}}\, d\omega\right| - \left|J_{\rho(t)} - \int_{\mathbb{S}^{d-1}} \omega \cdot M_{\Omega_{\rho(t)}}\, d\omega\right| \\ &\ge \left|\int_{\mathbb{S}^{d-1}} \omega \cdot M_{\Omega_{\rho(t)}}\, d\omega\right| - e^{-\frac{2(d-2)}{e^2}t}\Big(\int_{\mathbb{S}^{d-1}} \rho_0 \log \rho_0\, d\omega + 1 - \log C_M\Big) \\ &=: R(t).\end{aligned}$$

By (5.1), $C(d) := \left|\int_{\mathbb{S}^{d-1}} \omega \cdot M_{\Omega_\rho}\, d\omega\right|$ is a positive constant independent of Ω_ρ, thus

$$R(t) \to C(d) \quad \text{as } t \to \infty.$$

Recalling (3.27), this completes the proof. □

4.1 Proof of Theorem 2.2

We begin with

$$\begin{aligned}\frac{d}{dt} J_\rho &= \int \omega\, \nabla_\omega \cdot \big(\rho \nabla_\omega (\log \rho - \nabla_\omega(\omega \cdot \Omega_\rho))\big)\, d\omega \\ &= -\int \rho\, \nabla_\omega \log \rho\, d\omega + \int \rho\, \nabla_\omega(\omega \cdot \Omega_\rho)\, d\omega.\end{aligned}$$

If we regard the above terms as functionals on ρ, that is,

$$\begin{aligned}\mathcal{J}_1(\rho) &:= \int \rho\, \nabla_\omega \log \rho\, d\omega, \\ \mathcal{J}_2(\rho) &:= \int \rho\, \nabla_\omega(\omega \cdot \Omega_\rho)\, d\omega,\end{aligned}$$

then we see that

$$\mathcal{J}_1(M_{\Omega_\rho}) = \mathcal{J}_2(M_{\Omega_\rho}).$$

Also, noticing that $\mathcal{J}_1(\rho)$ and $\mathcal{J}_1(\rho)$ can be written as (see Section 2.2)

$$\begin{aligned}\mathcal{J}_1(\rho) &= \int \nabla_\omega \rho \, d\omega = (d-1) \int \omega \, \rho \, d\omega, \\ \mathcal{J}_2(\rho) &= \int \rho \, \mathbb{P}_{\omega^\perp} \Omega_\rho \, d\omega,\end{aligned}$$

we have

$$\begin{aligned}|\mathcal{J}_1(\rho) - \mathcal{J}_1(M_{\Omega_\rho})| &\leq (d-1) \Big| \int \omega \, (\rho - M_{\Omega_\rho}) \, d\omega \Big| \\ &\leq C \|\rho - M_{\Omega_\rho}\|_{L^1(\mathbb{S}^{d-1})},\end{aligned}$$

and

$$|\mathcal{J}_2(\rho) - \mathcal{J}_2(M_{\Omega_\rho})| \leq C \|\rho - M_{\Omega_\rho}\|_{L^1(\mathbb{S}^{d-1})}$$

for some dimensional constant C. Thus, thanks to Proposition 4.1, we have

$$\begin{aligned}\Big| \frac{d}{dt} J_\rho \Big| &= \Big| - \mathcal{J}_1(\rho) + \mathcal{J}_1(M_{\Omega_\rho}) - \mathcal{J}_2(M_{\Omega_\rho}) + \mathcal{J}_2(\rho) \Big| \\ &\leq C \|\rho - M_{\Omega_\rho}\|_{L^1(\mathbb{S}^{d-1})}. \\ &\leq C \, e^{-\frac{2(d-2)}{e^2} t}.\end{aligned}$$

Together with Lemma 4.2, this implies that there exist a nonzero constant vector $J_\infty \in \mathbb{R}^d$ such that

$$|J_{\rho(t)} - J_\infty| \leq C e^{-\frac{2(d-2)}{e^2} t}.$$

Therefore, setting $\Omega_\infty = \frac{J_\infty}{|J_\infty|}$, we have

$$\|M_{\Omega_{\rho(t)}} - M_{\Omega_\infty}\|_{L^1(\mathbb{S}^{d-1})} \leq C \, |\Omega_{\rho(t)} - \Omega_\infty| \leq C \, e^{-\frac{2(d-2)}{e^2} t},$$

that combined with Proposition 4.1 completes the proof. □

5 Uniqueness and Stability

In this section, we present a stability estimate in Wasserstein distance, which provides as a corollary the uniqueness result in Theorem 2.1. First of all, notice that the stability estimate (3.22) does not imply the stability for (1.1), because of the dependence on ε in (3.22).

We obtain here a stability estimate for short time when the two initial data $\rho_0, \bar{\rho}_0$ are close to each other as

$$W_2(\rho_0, \bar{\rho}_0) \leq \frac{|J_{\rho_0}|^2}{16}. \tag{5.1}$$

To get the stability estimate, we use the following lemma on the continuity of the momentum J_ρ with respect to the density ρ.

Lemma 5.1. *Let $\rho, \bar{\rho} \in \mathcal{P}(\mathbb{S}^{d-1})$ be any measures satisfying $|J_{\bar{\rho}}| > 0$. Then,*

$$\big||J_{\bar{\rho}}| - |J_\rho|\big| \leq \frac{2\, W_2(\bar{\rho}, \rho)}{|J_{\bar{\rho}}|}.$$

Proof. We follow the same arguments in the proof of Proposition 3.2. Let φ_0 be a $d^2/2$-convex function such that $\exp_\omega(\nabla_\omega \varphi_0)$ is the optimal map sending $\rho\, d\omega$ onto $\bar{\rho}\, d\omega$, and consider the unique geodesic $r \mapsto \alpha_r\, d\omega$ connecting $\alpha_0 = \rho$ to $\alpha_1 = \bar{\rho}$.

Similarly for each $r \in [0,1]$, let φ_r be a $d^2/2$-convex function such that $\exp_\omega(\nabla_\omega \varphi_r)$ is the optimal map sending $\alpha_r\, d\omega$ onto $\bar{\rho}\, d\omega$, and consider the geodesic $s \mapsto \alpha_{r,s}\, d\omega$ connecting $\alpha_{r,0} = \alpha_r$ to $\alpha_{r,1} = \bar{\rho}$. Notice that $s \mapsto \alpha_{r,s}\, d\omega$ satisfies the continuity equation in the sense of distributions:

$$\left.\frac{\partial}{\partial s}\right|_{s=0} \alpha_{r,s} = -\nabla_\omega \cdot (\alpha_{r,s} \nabla_\omega \varphi_r).$$

Using the same computations as in (3.24), we have

$$\frac{\partial}{\partial h}\bigg|_{h=r}\alpha_h = \frac{1}{1-r}\frac{\partial}{\partial s}\bigg|_{s=0}\alpha_{r,s} = -\frac{1}{1-r}\nabla_\omega\cdot(\alpha_r\nabla_\omega\varphi_r),$$

thus

$$\begin{aligned}
\frac{d}{dt}\bigg|_{h=r}|J(\alpha_h)|^2 &= 2J(\alpha_h)\cdot\frac{d}{dt}\bigg|_{h=r}J(\alpha_h)\\
&= \frac{2}{1-r}\int_{\mathbb{S}^{d-1}}\nabla_\omega(\omega\cdot J(\alpha_r))\nabla_\omega\varphi_r\alpha_r\,d\omega\\
&\leq \frac{2}{1-r}\sqrt{\int_{S^{d-1}}|\nabla_\omega(\omega\cdot J(\alpha_r))|^2\,\alpha_r\,d\omega}\sqrt{\int_{S^{d-1}}|\nabla_\omega\varphi_r|^2\,\alpha_r\,d\omega}\\
&\leq \frac{2}{1-r}W_2(\alpha_r,\bar\rho)\\
&\leq 2\,W_2(\rho,\bar\rho).
\end{aligned}$$

Integrating the above inequality from $r=0$ to $r=1$ we get

$$|J_{\bar\rho}|^2 - |J_\rho|^2 \leq 2\,W_2(\rho,\bar\rho).$$

Similarly applying the above arguments to another $d^2/2$-convex function $\overline{\varphi}_0$ sending $\overline{\rho}\,d\omega$ onto $\rho\,d\omega$ we get

$$|J_\rho|^2 - |J_{\bar\rho}|^2 \leq 2\,W_2(\rho,\bar\rho),$$

hence

$$\Big||J_{\bar\rho}| - |J_\rho|\Big| \leq \frac{2\,W_2(\bar\rho,\rho)}{|J_{\bar\rho}|+|J_\rho|} \leq \frac{2\,W_2(\bar\rho,\rho)}{|J_{\bar\rho}|}.$$

□

5.1 Proof of Theorem 2.3

Since ρ solves the continuity equation

$$\partial_t\rho + \nabla_\omega\cdot(\rho\,\nabla_\omega(\omega\cdot\Omega_\rho - \log\rho)) = 0,$$

it follows from the Benamou and Brenier formula [11] that, for any $t > 0$,

$$W_2^2(\rho(t), \rho_0) \leq t \int_0^t \int_{\mathbb{S}^{d-1}} |\nabla_\omega(\omega \cdot \Omega_{\rho(\tau)} - \log \rho(\tau))|^2 \rho(\tau)\, d\omega\, d\tau.$$

In addition, since

$$\frac{d}{dt}\mathcal{E}^0(\rho) = -\int_{\mathbb{S}^{d-1}} |\nabla_\omega(\log \rho - \omega \cdot \Omega_\rho)|^2 \rho\, d\omega,$$

we have

$$W_2^2(\rho(t), \rho_0) \leq t \int_0^t \Big(-\frac{d}{d\tau}\mathcal{E}^0(\rho)\Big)\, d\tau.$$

Recalling (4.1), we see that

$$\frac{d}{dt}\mathcal{E}^0(\rho) = \frac{d}{dt} H(\rho \mid M_{\Omega_\rho}).$$

Thus, for all $t > 0$,

$$\begin{aligned} &W_2^2(\rho(t), \rho_0) \\ &\qquad \leq t\Big(H(\rho_0 \mid M_{\Omega(\rho_0)}) - H(\rho(t) \mid M_{\Omega_{\rho(t)}})\Big) \leq t\, H(\rho_0 \mid M_{\Omega(\rho_0)}). \end{aligned}$$

Analogously

$$W_2^2(\bar\rho(t), \bar\rho_0) \leq t\, H(\bar\rho_0 \mid M_{\Omega(\bar\rho_0)}).$$

Therefore, setting

$$\delta := \frac{|J_{\rho_0}|^4}{2^8 \max\{H(\rho_0 \mid M_{\Omega(\rho_0)}), H(\bar\rho_0 \mid M_{\Omega(\bar\rho_0)})\}}$$

we have that, for all $t \leq \delta$,

$$W_2(\rho(t), \rho_0) \leq \frac{|J_{\rho_0}|^2}{16}, \qquad W_2(\bar\rho(t), \bar\rho_0) \leq \frac{|J_{\rho_0}|^2}{16}. \tag{5.2}$$

For each time $t \leq \delta$, we consider the unique geodesic $r \mapsto \alpha_r \, d\omega$ connecting $\alpha_0 = \rho(t)$ to $\alpha_1 = \bar{\rho}(t)$.

Then, using (5.1) and (5.2), we have

$$\begin{aligned} W_2(\rho_0, \alpha_r) &\leq W_2(\rho_0, \rho(t)) + W_2(\rho(t), \alpha_r) \\ &\leq W_2(\rho_0, \rho(t)) + W_2(\rho(t), \bar{\rho}(t)) \\ &\leq 2\, W_2(\rho_0, \rho(t)) + W_2(\rho_0, \bar{\rho}(t)) \\ &\leq 2\, W_2(\rho_0, \rho(t)) + W_2(\rho_0, \bar{\rho}_0) + W_2(\bar{\rho}_0, \bar{\rho}(t)) \\ &\leq \frac{|J_{\rho_0}|^2}{4}, \end{aligned}$$

and applying Lemma 5.1 to ρ_0 and α_r we get

$$\big| |J_{\rho_0}| - |J_{\alpha_r}| \big| \leq \frac{2\, W_2(\rho_0, \alpha_r)}{|J_{\rho_0}|} \leq \frac{|J_{\rho_0}|}{2},$$

thus

$$|J_{\alpha_r}| \geq \frac{|J_{\rho_0}|}{2}. \tag{5.3}$$

We now compute the second derivative of $\mathcal{E}^0$ using (3.24) and (3.2) with $\varepsilon = 0$, and thanks to (5.3) we have

$$\left.\frac{d^2}{dh^2}\right|_{h=r} \mathcal{E}^0(\alpha_h) \geq -\lambda\, W_2^2(\rho^\varepsilon(t), \bar{\rho}^\varepsilon(t)),$$

where $\lambda := (1 + 2/|J_{\rho_0}|) - (d-2)$.

Hence, using the same arguments in the proof of Proposition 3.2, we deduce that

$$W_2(\rho(t), \bar{\rho}(t)) \leq e^{\lambda t} W_2(\rho_0, \bar{\rho}_0) \qquad \forall\, t \in [0, \delta], \tag{5.4}$$

as desired. □

5.2 Proof of the uniqueness in Theorem 2.1

The short time stability estimate (5.4) implies the uniqueness of weak solutions to (1.1). Indeed, if $W_2(\rho_0, \bar{\rho}_0) = 0$, then $W_2(\rho(t), \bar{\rho}(t)) = 0$ for all $t \leq \delta$. Thanks to Lemma 4.2 and

$$\frac{d}{dt} H(\rho \mid M_{\Omega_\rho}) \leq 0,$$

a continuation argument implies $W_2(\rho(t), \bar{\rho}(t)) = 0$ for all $t \geq 0$. $\square$

Appendix A

We here present how to compute explicitely the momentum J_{M_Ω} of the Fisher-von Mises distribution in the case $d = 3$.

Let us fix a reference Cartesian coordinate system with $e_3 = \Omega$, and then consider the spherical coordinate system (θ, ϕ) associated with the orthonormal basis (e_1, e_2, Ω). Then a straightforward compution yields

$$C_M^{-1} = \int_{\mathbb{S}^2} e^{\omega \cdot \Omega} \, d\omega = \int_0^{2\pi} d\phi \int_0^{\pi} e^{\cos\theta} \sin\theta \, d\theta = 2\pi(e - e^{-1}).$$

Moreover, since

$$\omega = \sin\theta \cos\phi \, e_1 + \sin\theta \sin\phi \, e_2 + \cos\theta \, \Omega,$$

we have

$$J_{M_\Omega} = \int_{\mathbb{S}^2} \omega \, M_\Omega(\omega) \, d\omega = C_M \Omega \int_0^{2\pi} d\phi \int_0^{\pi} \cos\theta \, e^{\cos\theta} \sin\theta \, d\theta = \frac{2e^{-1}}{e - e^{-1}} \, \Omega.$$

Similarly using the generalized spherical coordinate system on $\mathbb{S}^{d-1}$, we have

$$J_{M_\Omega} = \frac{\int_0^{\pi} \cos\theta \, e^{\cos\theta} \sin^{d-2}\theta \, d\theta}{\int_0^{\pi} e^{\cos\theta} \sin^{d-2}\theta \, d\theta} \, \Omega. \tag{5.1}$$

Notice that C_M and $|J_{M_\Omega}|$ are constants only depending on dimension d, but independent of Ω.

Chapter 4

Dynamics of a spatially homogeneous Vicsek model for oriented particles on the plane

We consider a spatially homogeneous Kolmogorov-Vicsek model in two dimensions, which describes the alignment dynamics of self-driven stochastic particles that move on the plane at a constant speed, under space-homogeneity. In [44], Alessio Figalli and the authors have shown the existence of global weak solutions for this two-dimensional model. However, no time-asymptotic behavior has been obtained for the two-dimensional case, due to the failure of the celebrated Bakry-Emery condition for the logarithmic Sobolev inequality. We prove exponential convergence (with quantitative rate) of the weak solutions towards a Fisher-von Mises distribution, using a new condition for the logarithmic Sobolev inequality.

1 Introduction and Main results

Recently, the stochastic Vicsek model has received extensive attention in mathematical topics such as mean-field limits, hydrodynamic limits, and phase transitions [3, 16, 26, 29, 28, 31, 47, 50]. In this article, we study a time-asymptotic behavior of the so-called Kolmogorov-Vicsek model in two space dimensions. Such a model is governed by a nonlinear, nonlocal Fokker-Planck equation, which describes self-propelled stochastic particles moving on a plane with unit speed:

$$\begin{aligned} &\partial_t \rho = \Delta_\omega \rho - \nabla_\omega \cdot \Big(\rho\, \mathbb{P}_{\omega^\perp} \Omega_\rho\Big), \\ &\Omega_\rho = \frac{J_\rho}{|J_\rho|}, \quad J_\rho = \int_{\mathbb{S}^1} \omega\, \rho\, d\omega, \end{aligned} \tag{1.1}$$

where $\rho(t,\omega)$ is a probability density function at time t with direction $\omega \in \mathbb{S}^1$ (unit circle of $\mathbb{R}^2$), and the operators ∇_ω and Δ_ω denote the gradient and the Laplace-Beltrami operator on the circle $\mathbb{S}^1$. The force field $\mathbb{P}_{\omega^\perp}\Omega_\rho$ denotes the projection of the unit vector Ω_ρ onto the normal plane to ω, i.e., $\mathbb{P}_{\omega^\perp}\Omega_\rho := (Id - \omega \otimes \omega)\Omega_\rho$, which describes the mean-field force that governs the orientational interaction of self-driven particles by aligning them with the direction Ω_ρ determined by the flux J_ρ.

This model (1.1) is a spatially homogeneous version of the kinetic Kolmogorov-Vicsek model [31, 49], which was formally derived by Degond and Motsch [31] as a mean-field limit of the discrete Vicsek model [3, 26, 50, 15] with stochastic dynamics. Bolley, Cañizo and Carrillo [16] have rigorously justified the mean-field limit when the unit vector Ω_ρ in the force term of (1.1) is replaced by a more regular vector-field. As a study on phase transition, Degond, Frouvelle and Liu [28] provided a complete and rigorous description of phase transitions when Ω_ρ is replaced by $\nu(|J_\rho|)\Omega_\rho$, and there is a noise intensity $\tau(|J_\rho|)$ in front of $\Delta_\omega\rho$, where the functions ν and τ are Lipschitz, bounded, and satisfy that $|J_\rho| \mapsto \nu(|J_\rho|)/|J_\rho|$ and $|J_\rho| \mapsto \tau(|J_\rho|)$. It turns out that their modification leads to the appearance of phase transitions such as the number and nature of equilibria, stability, convergence rate, phase diagram and hysteresis, which depend on the ratio between ν and τ. We see that the assumptions of ν remove the singularity of Ω_ρ because $\nu(|J_\rho|)\Omega_\rho \to 0$ as $|J_\rho| \to 0$. This phase transition problem has been studied as well in [3, 26, 29, 28, 47, 50].

As a result on well-posedness of kinetic Vicsek model, Gamba and the first author [49] recently proved the existence and uniqueness of weak solutions to the spatially inhomogeneous Vicsek model under the a priori assumption of positivity of momentum, without handling the stability issues, whose difficulty is mainly coming from the facts that the momentum is not conserved and no dissipative energy functional has been found for the model. For a study for its numerical scheme, we refer to [48]. Frouvelle and Liu [47] have shown the well-posedness in the spatially homogeneous case with a more regular vector-field, i.e., $\mathbb{P}_{\omega^\perp} J_\rho$ instead of $\mathbb{P}_{\omega^\perp}\Omega_\rho$. Concerning studies on hydrodynamic descriptions of kinetic Vicsek model, we refer to [29, 28, 31, 32, 33, 46]; see also [17, 27, 30, 52] for other related studies.

In [44], the authors have shown the global-in-time existence of weak solutions to the two-dimensional model (1.1), and short-time stability in 2-

Wasserstein distance, whereas they have proved the convergence of the weak solutions towards Fisher-von Mises distribution in the higher-dimensional case for (1.1), that is, the space dimension is bigger than two. In order to show the exponential convergence to steady state, they have used the logarithmic Sobolev inequality based on the celebrated criterion of Bakry and Emery [10] (see Section 2), which requires the constraint on the space dimension.

In this article, we prove that the weak solution of (1.1) exponentially converges towards the Fisher-von Mises distribution as a stationary state. Notice that since $\mathbb{P}_{\omega^\perp}\Omega_\rho = \nabla_\omega(\omega \cdot \Omega_\rho)$, the equation (1.1) can be rewritten in the form:

$$\partial_t \rho = \nabla_\omega \cdot \Big(\rho \, \nabla_\omega \left(\log \rho - \omega \cdot \Omega_\rho\right)\Big), \tag{1.2}$$

which can be regarded as a gradient flow with respect to the Wasserstein distance of the free energy functional

$$\mathcal{E}(\rho) = \int_{\mathbb{S}^1} \rho \log \rho \, d\omega - |J_\rho|. \tag{1.3}$$

We can easily see that the equilibrium states of (1.2) have the form of the Fisher-von Mises distributions: for any given $\Omega \in \mathbb{S}^1$, these are given by

$$M_\Omega(\omega) := C_M e^{\omega \cdot \Omega},$$

where C_M is the following positive constant

$$C_M = \frac{1}{\int_{\mathbb{S}^1} e^{\omega \cdot \Omega} \, d\omega}, \tag{1.4}$$

so that M_Ω is a probability density function in $\mathcal{P}(\mathbb{S}^1)$; the space of probability measures in $\mathbb{S}^1$.

As in [44], we will show the time-asymptotic behavior using the relative entropy with respect to the Fisher-von Mises distribution M_{Ω_ρ} defined by

$$H(\rho|M_{\Omega_\rho}) := \int_{\mathbb{S}^1} \log\Big(\frac{\rho}{M_{\Omega_\rho}}\Big)\rho\, d\omega,$$

which actually control the L^1-distance between ρ and M_{Ω_ρ}. We show that $H(\rho|M_{\Omega_\rho})$ decays exponentially. The proof of such a decay relies on two main estimates. The first one is a localized version of the logarithmic Sobolev inequality which we prove in section 2. The second one is a growth control of the dissipation given in Lemma 3.1. The main heuristic idea for such control is coming from the finite dimensional identity

$$\frac{d}{dt}\langle \nabla_{\gamma(t)}E, \nabla_{\gamma(t)}E\rangle = -2\langle \nabla^2_{\gamma(t)}E\nabla_{\gamma(t)}E, \nabla_{\gamma(t)}E\rangle, \quad \forall t \geq 0,$$

which holds for any E in $\mathcal{C}^2(\mathbb{R}^n)$ and any γ in $\mathcal{C}^1([0,\infty), \mathbb{R}^n)$ satisfying $\dot\gamma = -\nabla_\gamma E$.

Indeed, in the setting of the formal Calculus introduced by Otto and Villani (see [84, Section 3]), such a identity corresponds to the connection between (3.5) and the formal expression for the Wasserstein Hessian of the free energy (1.3) appeared in [44, Lemma 3.1].

We now state the main results of the paper.

Theorem 1.1. *Let $\rho_0 \in \mathcal{P}(\mathbb{S}^1)$ be an initial probability measure satisfying*

$$\rho_0 > 0, \quad |J_{\rho_0}| > 0, \quad \int_{\mathbb{S}^1} \rho_0 \log \rho_0\, d\omega < \infty. \tag{1.5}$$

Then, the equation (1.1) *has a unique weak solution $\rho \in L^2_{loc}([0,\infty), W^{1,1}(\mathbb{S}^1))$, which is weakly continuous in time, and satisfies time-asymptotic behaviors as* *follows:*

1) There exist a constant unit vector $\Omega_\infty \in \mathbb{S}^1$ such that for all $t > 0$,

$$\|\rho_t - M_{\Omega_\infty}\|_{L^1(\mathbb{S}^1)} \lesssim H(\rho_0|M_{\Omega_{\rho_0}})e^{-Bt}. \tag{1.6}$$

Here, the exponential decay rate B depends on initial data as

$$B = \left(\frac{|J_{\rho_0}|e^{-T_0}}{2|J_{\rho_0}|e^{-T_0}+2}\left(\exp\left[\left(2+\frac{2}{|J_{\rho_0}|e^{-T_0}}\right)T_0\right]-1\right)\right.$$
$$\left.+ C_*\exp\left[\left(2+\frac{2}{|J_{\rho_0}|e^{-T_0}}\right)T_0\right]\right)^{-1},$$

where $T_0 > 0$ is some constant depending on initial data with $T_0 \lesssim H(\rho_0|M_{\Omega_{\rho_0}})$, and

$$C_* := \frac{2\pi^2 e^{2(1+|\log C_M|)}(1+\frac{1}{15}\varepsilon_*)}{1-\frac{7}{6}\varepsilon_*},$$

for any fixed positive constant $\varepsilon_* \le \frac{1}{10}$.

2) For any $t \ge T_0$, we have

$$\|\rho_t - M_{\Omega_\infty}\|_{L^1(\mathbb{S}^1)} \lesssim H(\rho_0|M_{\Omega_{\rho_0}})e^{-C_*^{-1}(t-T_0)}.$$

3) For any $\varepsilon > 0$, there exists $\delta > 0$ such that if $H(\rho_0|M_{\Omega_{\rho_0}}) < \delta$, then for all $t > 0$,

$$\|\rho_t - M_{\Omega_\infty}\|_{L^1(\mathbb{S}^1)} \lesssim H(\rho_0|M_{\Omega_{\rho_0}})\exp\left[-\left(\frac{1}{2\pi^2 e^{2(1+|\log C_M|)}}-\varepsilon\right)t\right].$$

Remark 1.1. *The estimates (1.6) in Theorem (1.1) represent the exponential convergence of weak solutions towards some steady state M_{Ω_∞}, where it is not clear how to determine the vector Ω_∞ from the initial data ρ_0, because the momentum J_{ρ_t} is not conserved in time.*

Remark 1.2. *Following the proof in Section 3, the estimates* (1.6) *under the above three conditions 1)-3) are straightforward results of the more detailed estimates (see* (3.17)*):*

$$\|\rho_t - M_{\Omega_\infty}\|_{L^1(\mathbb{S}^1)} \le H(\rho_0|M_{\Omega_{\rho_0}})\left(e^{-\int_0^t B(s)ds} + C\int_t^\infty e^{-\int_0^r B(s)ds}dr\right), \quad \forall t > 0, \tag{1.7}$$

where $C > 0$ is some constant and $B(t)$ is a positive function defined by (with notation $f_+ := \max(f, 0)$)

$$\begin{aligned} B(t) := \Bigg(\frac{|J_{\rho_0}|e^{-T_0}}{2|J_{\rho_0}|e^{-T_0} + 2} \Bigg(\exp\left[\left(2 + \frac{2}{|J_{\rho_0}|e^{-T_0}}\right)T_0\right] \\ - \exp\left[\left(2 + \frac{2}{|J_{\rho_0}|e^{-T_0}}\right)t\right]\Bigg)_+ \\ + C_* \exp\left[\left(2 + \frac{2}{|J_{\rho_0}|e^{-T_0}}\right)(T_0 - t)_+\right]\Bigg)^{-1}. \end{aligned}$$

Here, $T_0 > 0$ is some constant such that

$$T_0 \le 2H(\rho_0|M_{\Omega_{\rho_0}})\Big[\min\Big(C_M^2 e^{-2}\varepsilon_*^2, \frac{L}{2C_*}C_M^2 e^{-2}\varepsilon_*^2, mC_*^{-1}\Big)\Big]^{-1},$$

where C_*, ε_* are the constants appeared in Theorem 1.1, and

$$L := \left(2 + \frac{4}{m}\right)^{-1} \log 2,$$

where m denotes a strength of momentum of the Fisher-von Mises distribution, i.e.,

$$m := \Big| \int_{\mathbb{S}^1} \omega M_{\Omega_\rho}\, d\omega \Big|,$$

which is a constant (see [44, Appendix]).

Remark 1.3. *In the proof of Theorem 1.1, we first obtain the exponential convergence of* $H(\rho_t|M_{\Omega_t})$ *as*

$$H(\rho_t|M_{\Omega_{\rho_t}}) \leq H(\rho_0|M_{\Omega_{\rho_0}})e^{-\int_0^t B(s)ds}. \tag{1.8}$$

However, the Fisher-von Mises distribution $M_{\Omega_{\rho_t}}$ *is not constant in time, because of no conservation of momentum* J_{ρ_t}*. In fact, we show that* Ω_{ρ_t} *stabilizes to a fixed vector as* $t \to \infty$*, observing that* $|\frac{d}{dt}J_{\rho_t}|$ *vanishes as* $t \to \infty$ *thanks to the decay estimate* (1.8).

Remark 1.4. *Notice that since*

$$\int_{\mathbb{S}^1} \rho_0 \log \rho_0 \, d\omega - 1 - \log C_M \leq H(\rho_0 \mid M_{\Omega_{\rho_0}}) \leq \int_{\mathbb{S}^1} \rho_0 \log \rho_0 \, d\omega + 1 - \log C_M,$$

we see that $H(\rho_0 \mid M_{\Omega_{\rho_0}}) < \infty$ *if and only if* $\int_{\mathbb{S}^1} \rho_0 \log \rho_0 \, d\omega < \infty$*. Therefore, the initial condition* $\int_{\mathbb{S}^1} \rho_0 \log \rho_0 \, d\omega < \infty$ *in* (1.5) *can be replaced by* $H(\rho_0 \mid M_{\Omega_{\rho_0}}) < \infty$.

In [44], the authors have proven that for any dimension $n \geq 1$, the model (1.1) on n-dimensional sphere $\mathbb{S}^n$ has a weak solution $\rho \in L^2_{loc}([0,\infty), W^{1,1}(\mathbb{S}^n))$ that satisfies

$$|J_{\rho_t}| \geq |J_{\rho_0}|e^{-nt}, \quad \forall t > 0. \tag{1.9}$$

Additionally, we showed short-time stability in 2-Wasserstein distance W_2 as follows: for any probability measure ρ_0 and $\bar{\rho}_0$ satisfying (1.5) and

$$W_2(\rho_0, \bar{\rho}_0) \leq \frac{|J_{\rho_0}|}{16},$$

there exists $\delta > 0$ such that any two solutions ρ_t and $\bar{\rho}_t$ of (1.1) starting from ρ_0 and $\bar{\rho}_0$ satisfies

$$W_2(\rho_t, \bar{\rho}_t) \leq \exp\left[\left(2 - n + \frac{2}{|J_{\rho_0}|}\right)t\right] W_2(\rho_0, \bar{\rho}_0), \quad \forall t < \delta, \tag{1.10}$$

where the short time δ is explicitly found as

$$\delta = \frac{|J_{\rho_0}|^4}{2^8 \max\{H(\rho_0|M_{\Omega_{\rho_0}}), H(\bar{\rho}_0|M_{\Omega_{\bar{\rho}_0}})\}}. \tag{1.11}$$

In fact, for this stability estimate, the authors have not used the logarithmic Sobolev inequality based on the criterion of Bakry and Emery. Therefore, the above stability still holds when $n = 1$, whose case will be added in [44]. On the other hand, the uniqueness of the weak solutions has been proven in the case of $n \geq 2$, by using the stability estimate (1.10) up to δ in (1.11) together with the exponential convergence towards Fisher-von Mises distribution.

Therefore, once we get the exponential convergence (3.17) in Theorem 1.1, the uniqueness of the weak solution to (1.1) holds true by the same argument in [44] as mentioned above. Recently, we recognized that we can prove the uniqueness, based on (1.10) and the energy method with the parabolic regularity without using the large-time behavior. We give its proof in Appendix B as an another proof of the uniqueness.

The paper is organized as follows. In the next section, we present a localized version for logarithmic Sobolev inequality, which is crucially used in the proof of Theorem 1.1 in Section 3.

2 Logarithmic Sobolev inequality

In this section, we present a simple condition for the logarithmic Sobolev inequality on $\mathbb{S}^1$ equipped with the ambient metric. Such condition is one of the crucial ingredients that we use to show the exponential decay (3.17). We first fix a reference probability measure $e^{-\Psi(\omega)}d\omega$ such that $\Psi \in C^1(\mathbb{S}^1)$. Then, for a given probability measure $\rho\, d\omega$, we define its relative entropy with respect to $e^{-\Psi(\omega)}d\omega$ by

$$H(\rho|e^{-\Psi}) = \int_{\mathbb{S}^1} \log\left(\frac{\rho}{e^{-\Psi}}\right)\rho\, d\omega,$$

and the relative Fisher information by

$$I(\rho|e^{-\Psi}) = \int_{\mathbb{S}^1} \left|\nabla \log \frac{\rho}{e^{-\Psi}}\right|^2 \rho\, d\omega,$$

where ∇ denotes the gradient on $\mathbb{S}^1$.

We recall the celebrated criterion of Bakry and Emery [10] for a logarithmic Sobolev inequality as follows: If there exists a constant $\alpha > 0$ such that

$$D^2\Psi + \text{Ricci curvature tensor on } \mathbb{S}^1 \geq \alpha I_n, \tag{2.1}$$

then the probability measure $e^{-\Psi}d\omega$ satisfies a logarithmic Sobolev inequality with α, i.e., for all probability measure $\rho\, d\omega$ we have

$$H(\rho|e^{-\Psi}) \leq \frac{1}{2\alpha} I(\rho|e^{-\Psi}).$$

In [44], we proved the exponential convergence estimate like (3.17) with explicit decay rate, using the above criterion (2.1) together with the facts that the Ricci curvature tensor of $\mathbb{S}^n$ is $(n-1)I_n$, and the logarithmic Sobolev inequality is stable under bounded perturbations (see [63, 84]). However, if we consider the 1-dimensional sphere $\mathbb{S}^1$, its Ricci curvature vanishes. Therefore, the condition (2.1) is not satisfied anymore when Ψ takes the form $\Psi(\omega) = -\omega \cdot \Omega$, as a Fisher-von Mises distribution. We will overcome this difficulty by imposing a smallness condition of the L^∞ distance between ρ and $e^{-\Psi}$ in the following Lemma.

Lemma 2.1. *Let ρ and $e^{-\Psi}$ be probability measures on $\mathbb{S}^1$ such that $\rho, \Psi \in C^1(\mathbb{S}^1)$, and $|\Psi| \le \lambda$, for some constant λ. For any fixed positive constant $\varepsilon_* \le \frac{1}{10}$, if*

$$\|\rho e^{\Psi} - 1\|_{L^\infty(\mathbb{S}^1)} \le \varepsilon_*,$$

then,

$$H(\rho|e^{-\Psi}) \le \frac{2\pi^2 e^{2\lambda}(1+\frac{1}{15}\varepsilon_*)}{1-\frac{7}{6}\varepsilon_*} I(\rho|e^{-\Psi}). \tag{2.2}$$

Proof. We set $f := \rho e^{\Psi} - 1$, $\gamma(s) := (1+sf)e^{-\Psi}$ for $s \in [0,1]$. Then we define

$$H(\gamma(s)|e^{-\Psi}) = \int_{\mathbb{S}^1} (1+sf)e^{-\Psi} \log(1+sf) d\omega,$$

and

$$I(\gamma(s)|e^{-\Psi}) = \int_{\mathbb{S}^1} (1+sf)e^{-\Psi} |\nabla \log(1+sf)|^2 d\omega.$$

for every s in$[0,1]$.

We will use the Taylor theorem and the Poincaré inequality to show (2.2). First of all, we see that

$$H(\gamma(0)|e^{-\Psi}) = 0 = I(\gamma(0)|e^{-\Psi}).$$

A straightforward computation yields

$$
\begin{aligned}
\frac{d}{ds}H(\gamma(s)|e^{-\Psi}) &= \int_{\mathbb{S}^1}(1+\log(1+sf))fe^{-\Psi}d\omega,\\
\frac{d^2}{ds^2}H(\gamma(s)|e^{-\Psi}) &= \int_{\mathbb{S}^1}\frac{f^2}{1+sf}e^{-\Psi}d\omega,\\
\frac{d^3}{ds^3}H(\gamma(s)|e^{-\Psi}) &= -\int_{\mathbb{S}^1}\frac{f^3}{(1+sf)^2}e^{-\Psi}d\omega.
\end{aligned}
$$

Setting $g(s)=1+sf$, by direct computation, we obtain

$$
\begin{aligned}
\frac{d}{ds}I(\gamma(s)|e^{-\Psi}) &= \int_{\mathbb{S}^1}\nabla\log(g(s))\Big[2\nabla\Big(\frac{g'(s)}{g(s)}\Big)g(s)+\nabla\log(g(s))g'(s)\Big]e^{-\Psi}d\omega,\\
\frac{d^2}{ds^2}I(\gamma(s)|e^{-\Psi}) &= \int_{\mathbb{S}^1}\nabla\Big(\frac{g'(s)}{g(s)}\Big)\Big[2\nabla\Big(\frac{g'(s)}{g(s)}\Big)g(s)+\nabla\log(g(s))g'(s)\Big]e^{-\Psi}d\omega\\
&\quad+\int_{\mathbb{S}^1}\nabla\log(g(s))\Big[-2\nabla\Big(\frac{|g'(s)|^2}{g(s)^2}\Big)g(s)+3\nabla\Big(\frac{g'(s)}{g(s)}\Big)g'(s)\Big]e^{-\Psi}d\omega,\\
\frac{d^3}{ds^3}I(\gamma(s)|e^{-\Psi}) &= -\int_{\mathbb{S}^1}\nabla\Big(\frac{|g'(s)|^2}{g(s)^2}\Big)\Big[2\nabla\Big(\frac{g'(s)}{g(s)}\Big)g(s)\\
&+\nabla\log(g(s))g'(s)\Big]e^{-\Psi}d\omega\\
&\quad+2\int_{\mathbb{S}^1}\nabla\Big(\frac{g'(s)}{g(s)}\Big)\Big[-2\nabla\Big(\frac{|g'(s)|^2}{g(s)^2}\Big)g(s)+3\nabla\Big(\frac{g'(s)}{g(s)}\Big)g'(s)\Big]e^{-\Psi}d\omega\\
&\quad+\int_{\mathbb{S}^1}\nabla\log(g(s))\Big[4\nabla\Big(\frac{|g'(s)|^3}{g(s)^3}\Big)g(s)-5\nabla\Big(\frac{|g'(s)|^2}{g(s)^2}\Big)g'(s)\Big]e^{-\Psi}d\omega,
\end{aligned}
$$

where we have used the fact that$g''(s)=0$.

Since

$$
\begin{aligned}
\frac{d}{ds}\Big|_{s=0}H(\gamma(s)|e^{-\Psi}) &= \int_{\mathbb{S}^1}fe^{-\Psi}d\omega=\int_{\mathbb{S}^1}(\rho-e^{-\Psi})d\omega=0,\\
\frac{d}{ds}\Big|_{s=0}I(\gamma(s)|e^{-\Psi}) &= 0,
\end{aligned}
$$

it follows from Taylor theorem that there exists $s_1, s_2 \in (0,1)$ such that

$$
\begin{aligned}
H(\rho|e^{-\Psi}) &= H(\gamma(1)|e^{-\Psi}) \\
&= \frac{1}{2}\int_{\mathbb{S}^1} f^2 e^{-\Psi} d\omega + \frac{1}{6}\frac{d^3}{ds^3}\Big|_{s=s_1} H(\gamma(s)|e^{-\Psi}), \\
I(\rho|e^{-\Psi}) &= I(\gamma(1)|e^{-\Psi}) \\
&= \int_{\mathbb{S}^1} |\nabla f|^2 e^{-\Psi} d\omega + \frac{1}{6}\frac{d^3}{ds^3}\Big|_{s=s_2} I(\gamma(s)|e^{-\Psi}).
\end{aligned}
\tag{2.3}
$$

Since $\int_{\mathbb{S}^1} f e^{-\Psi} d\omega = 0$ and $e^{-\Psi} > 0$, there exists $\omega_0 \in \mathbb{S}^1$ such that $f(\omega_0) = 0$. Then, using $|\Psi| \le \lambda$, we have a Poincaré inequality:

$$
\begin{aligned}
\int_{\mathbb{S}^1} f^2 e^{-\Psi} d\omega &\le e^{\lambda}\int_{\mathbb{S}^1} f^2 d\omega = e^{\lambda}\int_{\mathbb{S}^1} (f(\omega) - f(\omega_0))^2 d\omega = e^{\lambda}\int_{\mathbb{S}^1}\Big(\int_{\omega_0}^{\omega}\nabla f\Big)^2 d\omega \\
&\le 4e^{\lambda}\pi^2 \int_{\mathbb{S}^1} |\nabla f|^2 d\omega \le 4e^{2\lambda}\pi^2 \int_{\mathbb{S}^1} |\nabla f|^2 e^{-\Psi} d\omega.
\end{aligned}
\tag{2.4}
$$

A straightforward computation with $\|f\|_{L^\infty(\mathbb{S}^1)} \le \varepsilon_* \le \frac{1}{10}$ gives the estimates for the third order terms in (2.3) as follows:

$$
\Big|\frac{d^3}{ds^3}\Big|_{s=s_1} H(\gamma(s)|e^{-\Psi})\Big| \le \frac{1}{5}\varepsilon_* \int_{\mathbb{S}^1} f^2 e^{-\Psi} d\omega
$$

and

$$
\Big|\frac{d^3}{ds^3}\Big|_{s=s_2} I(\gamma(s)|e^{-\Psi})\Big| \le 42\varepsilon_* \int_{\mathbb{S}^1} |\nabla f|^2 e^{-\Psi} d\omega,
$$

Therefore, using (2.3), (2.4), and the above estimates we get

$$
\begin{aligned}
H(\rho|e^{-\Psi}) &\le \frac{1}{2}\Big(1 + \frac{1}{15}\varepsilon_*\Big)\int_{\mathbb{S}^1} f^2 e^{-\Psi} d\omega \\
&\le 2\pi^2 e^{2\lambda}\Big(1 + \frac{1}{15}\varepsilon_*\Big)\int_{\mathbb{S}^1} |\nabla f|^2 e^{-\Psi} d\omega \\
&\le \frac{2\pi^2 e^{2\lambda}\Big(1 + \frac{1}{15}\varepsilon_*\Big)}{1 - \frac{7}{6}\varepsilon_*} I(\rho|e^{-\Psi}).
\end{aligned}
$$

□

Remark 2.1. *In the proof of Lemma 2.1, we see that the Poincaré inequality implies the logarithmic Sobolev inequality* (2.2)*, provided* L^∞ *distance between* ρ *and* $e^{-\Psi}$ *is suitably small on* $\mathbb{S}^1$. *In general, we refer to [90] (see also [84]) on the implication of the Poincaré inequality as a linearization of the logarithmic Sobolev inequality, on a compact Riemannian manifold.*

3 Proof of Theorem 1.1

We begin by observing that a straightforward computation with (1.2) implies

$$\frac{d}{dt} H(\rho_t | M_{\Omega_{\rho_t}}) = -I(\rho_t | M_{\Omega_{\rho_t}}). \tag{3.1}$$

We first show that $H(\rho | M_{\Omega_\rho})$ exponentially decays after a suitably large time using (3.1) and Lemma 2.1. Notice that the weak solution ρ to (1.1) becomes smooth instantaneously thanks to the parabolic regularization, therefore we can consider smoothness of weak solutions ρ to (1.1) for the decay estimates. Moreover, since $\rho_0 > 0$ by the assumption, we see that the weak solution ρ of the parabolic equation (1.1) is positive for all time, i.e., $\rho_t > 0$ for all $t > 0$. Since $\mathbb{S}^1$ is one-dimensional manifold, we easily see that $\|\rho M_{\Omega_\rho}^{-1} - 1\|_{L^\infty(\mathbb{S}^1)}$ is controlled by the dissipation $I(\rho | M_{\Omega_\rho})$ as follows:

$$\|\rho M_{\Omega_\rho}^{-1} - 1\|_{L^\infty(\mathbb{S}^1)} \leq C_M^{-1} e \sqrt{I(\rho | M_{\Omega_\rho})}. \tag{3.2}$$

Indeed, since $\int_{\mathbb{S}^1} \rho d\omega = \int_{\mathbb{S}^1} M_{\Omega_\rho} d\omega = 1$, there exists $\omega_0 \in \mathbb{S}^1$ such that $\rho(\omega_0) = M_{\Omega_\rho}(\omega_0)$, which together with $C_M e^{-1} \le M_{\Omega_\rho} \le C_M e$ imply that

$$\begin{aligned}\rho M_{\Omega_\rho}^{-1}(\omega) - 1 = \rho M_{\Omega_\rho}^{-1}(\omega) - \rho M_{\Omega_\rho}^{-1}(\omega_0) &= \int_{\omega_0}^{\omega} \nabla(\rho M_{\Omega_\rho}^{-1})\, d\omega \\ &\le \sqrt{C_M^{-1} e}\Big(\int_{\mathbb{S}^1} \frac{|\nabla(\rho M_{\Omega_\rho}^{-1})|^2}{\rho M_{\Omega_\rho}^{-1}} d\omega \Big)^{1/2} \\ &\le C_M^{-1} e \Big(\int_{\mathbb{S}^1} \Big| \frac{\nabla(\rho M_{\Omega_\rho}^{-1})}{\rho M_{\Omega_\rho}^{-1}} \Big|^2 \rho\, d\omega \Big)^{1/2} \\ &= C_M^{-1} e \sqrt{I(\rho | M_{\Omega_\rho})}.\end{aligned}$$

We will show that for ε_* appeared in Lemma 2.1, there exists $T_0 > 0$ such that

$$I(\rho_t | M_{\Omega_{\rho_t}}) \le C_M^2 e^{-2} \varepsilon_*^2, \quad \forall t \ge T_0, \tag{3.3}$$

therefore we use Lemma 2.1 together with (3.1) to get the exponential decay of $H(\rho | M_{\Omega_\rho})$.

Before showing (3.3), we first prove the following lemma on growth estimates of the dissipation $I(\rho | M_{\Omega_\rho})$.

Lemma 3.1. *Let $\rho_0 \in \mathcal{P}(\mathbb{S}^1)$ be an initial probability measure satisfying* (1.5). *Then, the solution ρ of* (1.1) *starting from ρ_0 satisfies that for any $t \ge s > 0$,*

$$I(\rho_t | M_{\Omega_{\rho_t}}) \le I(\rho_s | M_{\Omega_{\rho_s}}) e^{C(t,s)(t-s)}, \tag{3.4}$$

where

$$C(t,s) = 2 + \frac{2}{\min_{r \in [s,t]} |J_{\rho_r}|}.$$

Proof. We claim that

$$
\begin{aligned}
\frac{d}{dt} I(\rho_t | M_{\Omega_{\rho_t}}) &= -2 \int_{\mathbb{S}^1} \mathrm{tr}\Big(\Big[D^2 \log \frac{\rho_t}{e^{\omega\cdot\Omega_{\rho_t}}} \Big]^T D^2 \log \frac{\rho_t}{e^{\omega\cdot\Omega_{\rho_t}}} \Big) \rho_t \, d\omega \\
&\quad + 2 \int_{\mathbb{S}^1} \nabla \log \frac{\rho_t}{e^{\omega\cdot\Omega_{\rho_t}}} D^2(\omega\cdot\Omega_{\rho_t}) \nabla \log \frac{\rho_t}{e^{\omega\cdot\Omega_{\rho_t}}} \rho_t \, d\omega \\
&\quad + \frac{2}{|J_{\rho_t}|} \Big(\Big| \int_{\mathbb{S}^1} \nabla \log \frac{\rho_t}{e^{\omega\cdot\Omega_{\rho_t}}} \rho_t d\omega \Big|^2 - \Big| \int_{\mathbb{S}^1} \Omega_{\rho_t} \cdot \nabla \log \frac{\rho_t}{e^{\omega\cdot\Omega_{\rho_t}}} \rho_t d\omega \Big|^2 \Big).
\end{aligned} \tag{3.5}
$$

Once we show the above equality, since the first term in r.h.s. of (3.5) is non-positive and the Hessian of the map $\omega \mapsto \omega \cdot \Omega_{\rho_t}$ has norm bounded by 1, we have

$$
\begin{aligned}
\frac{d}{dt} I(\rho_t | M_{\Omega_{\rho_t}}) &\le \Big(2 + \frac{2}{|J_{\rho_t}|} \Big) \int_{\mathbb{S}^1} \Big| \nabla \log \frac{\rho_t}{e^{\omega\cdot\Omega_{\rho_t}}} \Big|^2 \rho_t \, d\omega \\
&= \Big(2 + \frac{2}{|J_{\rho_t}|} \Big) I(\rho_t | M_{\Omega_{\rho_t}}),
\end{aligned}
$$

which provides the desired inequality (3.4). Therefore, it remains to prove the claim (3.5).

First of all, we separate $I(\rho_t|M_{\Omega_{\rho_t}})$ into two parts as

$$
\begin{aligned}
I(\rho_t | M_{\Omega_{\rho_t}}) &= \int_{\mathbb{S}^1} \nabla \log \frac{\rho_t}{e^{\omega\cdot\Omega_{\rho_t}}} \cdot \nabla \log \rho_t \rho_t \, d\omega \\
&- \int_{\mathbb{S}^1} \nabla \log \frac{\rho_t}{e^{\omega\cdot\Omega_{\rho_t}}} \cdot \nabla(\omega \cdot \Omega_{\rho_t}) \rho_t \, d\omega \\
&=: I_1 + I_2,
\end{aligned}
$$

Using the Eq. (1.2) and integration by parts, we compute $\frac{d}{dt}I_1$ as follows:

$$
\begin{aligned}
\frac{d}{dt}I_1 &= \int_{\mathbb{S}^1} \nabla\partial_t \log\frac{\rho_t}{e^{\omega\cdot\Omega_{\rho_t}}}\cdot\nabla\rho_t\,d\omega + \int_{\mathbb{S}^1}\nabla\log\frac{\rho_t}{e^{\omega\cdot\Omega_{\rho_t}}}\cdot\nabla\partial_t\rho_t\,d\omega\\
&= -\int_{\mathbb{S}^1}\partial_t\log\frac{\rho_t}{e^{\omega\cdot\Omega_{\rho_t}}}\Delta\rho_t\,d\omega - \int_{\mathbb{S}^1}\Delta\log\frac{\rho_t}{e^{\omega\cdot\Omega_{\rho_t}}}\nabla\cdot\Big(\nabla\log\frac{\rho_t}{e^{\omega\cdot\Omega_{\rho_t}}}\rho_t\Big)\,d\omega\\
&= -\int_{\mathbb{S}^1}\frac{e^{\omega\cdot\Omega_{\rho_t}}}{\rho_t}\partial_t\frac{\rho_t}{e^{\omega\cdot\Omega_{\rho_t}}}\Delta\rho_t\,d\omega + \int_{\mathbb{S}^1}\nabla\Delta\log\frac{\rho_t}{e^{\omega\cdot\Omega_{\rho_t}}}\cdot\nabla\log\frac{\rho_t}{e^{\omega\cdot\Omega_{\rho_t}}}\rho_t\,d\omega\\
&= -\int_{\mathbb{S}^1}\frac{\partial_t\rho_t}{\rho_t}\Delta\rho_t\,d\omega + \int\partial_t(\omega\cdot\Omega_{\rho_t})\Delta\rho_t\,d\omega\\
&+\int_{\mathbb{S}^1}\nabla\Delta\log\frac{\rho_t}{e^{\omega\cdot\Omega_{\rho_t}}}\cdot\nabla\log\frac{\rho_t}{e^{\omega\cdot\Omega_{\rho_t}}}\rho_t\,d\omega.
\end{aligned}
$$

Similarly, compute $\frac{d}{dt}I_2$ as follow:

$$
\begin{aligned}
\frac{d}{dt}I_2 &= -\int_{\mathbb{S}^1}\nabla\partial_t\log\frac{\rho_t}{e^{\omega\cdot\Omega_{\rho_t}}}\cdot\nabla(\omega\cdot\Omega_{\rho_t})\rho_t\,d\omega\\
&\quad-\int_{\mathbb{S}^1}\nabla\log\frac{\rho_t}{e^{\omega\cdot\Omega_{\rho_t}}}\cdot\nabla\partial_t(\omega\cdot\Omega_{\rho_t})\rho_t\,d\omega\\
&\quad-\int_{\mathbb{S}^1}\nabla\log\frac{\rho_t}{e^{\omega\cdot\Omega_{\rho_t}}}\cdot\nabla(\omega\cdot\Omega_{\rho_t})\partial_t\rho_t\,d\omega\\
&=: I_{21}+I_{22}+I_{23},
\end{aligned}
$$

where the three terms are computed as

$$
\begin{aligned}
I_{21} &= \int_{\mathbb{S}^1}\partial_t\log\frac{\rho_t}{e^{\omega\cdot\Omega_{\rho_t}}}\nabla\cdot\Big(\nabla(\omega\cdot\Omega_{\rho_t})\rho_t\Big)\,d\omega\\
&= \int_{\mathbb{S}^1}\Big(\frac{\partial_t\rho_t}{\rho_t}-\partial_t(\omega\cdot\Omega_{\rho_t})\Big)\Big(\Delta(\omega\cdot\Omega_{\rho_t})\rho_t+\nabla(\omega\cdot\Omega_{\rho_t})\cdot\nabla\rho_t\Big)\,d\omega,\\
I_{22} &= -\int_{\mathbb{S}^1}\Big(\nabla\rho_t-\nabla(\omega\cdot\Omega_{\rho_t})\rho_t\Big)\cdot\nabla\partial_t(\omega\cdot\Omega_{\rho_t})\,d\omega\\
&= \int_{\mathbb{S}^1}\Big(\Delta\rho_t-\Delta(\omega\cdot\Omega_{\rho_t})\rho_t-\nabla(\omega\cdot\Omega_{\rho_t})\cdot\nabla\rho_t\Big)\partial_t(\omega\cdot\Omega_{\rho_t})\,d\omega,\\
I_{23} &= -\int_{\mathbb{S}^1}\nabla\log\frac{\rho_t}{e^{\omega\cdot\Omega_{\rho_t}}}\cdot\nabla(\omega\cdot\Omega_{\rho_t})\nabla\cdot\Big(\nabla\log\frac{\rho_t}{e^{\omega\cdot\Omega_{\rho_t}}}\rho_t\Big)\,d\omega\\
&= \int_{\mathbb{S}^1}\nabla\Big(\nabla\log\frac{\rho_t}{e^{\omega\cdot\Omega_{\rho_t}}}\cdot\nabla(\omega\cdot\Omega_{\rho_t})\Big)\cdot\nabla\log\frac{\rho_t}{e^{\omega\cdot\Omega_{\rho_t}}}\rho_t\,d\omega
\end{aligned}
$$

Combining the above computations, we have

$$
\begin{aligned}
&\frac{d}{dt}I(\rho_t|M_{\Omega_{\rho_t}})\\
&\quad= \int_{\mathbb{S}^1}\biggl(\nabla\Delta\log\frac{\rho_t}{e^{\omega\cdot\Omega_{\rho_t}}}\cdot\nabla\log\frac{\rho_t}{e^{\omega\cdot\Omega_{\rho_t}}}\\
&+\nabla\Bigl(\nabla\log\frac{\rho_t}{e^{\omega\cdot\Omega_{\rho_t}}}\cdot\nabla(\omega\cdot\Omega_{\rho_t})\Bigr)\cdot\nabla\log\frac{\rho_t}{e^{\omega\cdot\Omega_{\rho_t}}}\biggr)\rho_t\,d\omega\\
&\qquad-\int_{\mathbb{S}^1}\biggl(\frac{\Delta\rho_t}{\rho_t}-\Delta(\omega\cdot\Omega_{\rho_t})-\frac{\nabla(\omega\cdot\Omega_{\rho_t})\cdot\nabla\rho_t}{\rho_t}\biggr)\partial_t\rho_t\,d\omega\\
&\qquad+2\int_{\mathbb{S}^1}\Bigl(\Delta\rho_t-\Delta(\omega\cdot\Omega_{\rho_t})\rho_t-\nabla(\omega\cdot\Omega_{\rho_t})\cdot\nabla\rho_t\Bigr)\partial_t(\omega\cdot\Omega_{\rho_t})\,d\omega\\
&\quad=: J_1+J_2+J_3.
\end{aligned}
$$

We first use (A.1) and (1.2) to get

$$
\begin{aligned}
J_2 &= -\int_{\mathbb{S}^1}\biggl(\Delta\log\frac{\rho_t}{e^{\omega\cdot\Omega_{\rho_t}}}+|\nabla\log\frac{\rho_t}{e^{\omega\cdot\Omega_{\rho_t}}}|^2+\nabla(\omega\cdot\Omega_{\rho_t})\cdot\nabla\log\frac{\rho_t}{e^{\omega\cdot\Omega_{\rho_t}}}\biggr)\partial_t\rho_t\,d\omega\\
&= \int_{\mathbb{S}^1}\nabla\biggl(\Delta\log\frac{\rho_t}{e^{\omega\cdot\Omega_{\rho_t}}}+|\nabla\log\frac{\rho_t}{e^{\omega\cdot\Omega_{\rho_t}}}|^2\\
&+\nabla(\omega\cdot\Omega_{\rho_t})\cdot\nabla\log\frac{\rho_t}{e^{\omega\cdot\Omega_{\rho_t}}}\biggr)\cdot\nabla\log\frac{\rho_t}{e^{\omega\cdot\Omega_{\rho_t}}}\rho_t\,d\omega\\
&= \int_{\mathbb{S}^1}\nabla\Delta\log\frac{\rho_t}{e^{\omega\cdot\Omega_{\rho_t}}}\cdot\nabla\log\frac{\rho_t}{e^{\omega\cdot\Omega_{\rho_t}}}\rho_t\,d\omega\\
&+\int_{\mathbb{S}^1}\nabla|\nabla\log\frac{\rho_t}{e^{\omega\cdot\Omega_{\rho_t}}}|^2\cdot\nabla\log\frac{\rho_t}{e^{\omega\cdot\Omega_{\rho_t}}}\rho_t\,d\omega\\
&\quad+\int_{\mathbb{S}^1}\nabla\Bigl(\nabla(\omega\cdot\Omega_{\rho_t})\cdot\nabla\log\frac{\rho_t}{e^{\omega\cdot\Omega_{\rho_t}}}\Bigr)\cdot\nabla\log\frac{\rho_t}{e^{\omega\cdot\Omega_{\rho_t}}}\rho_t\,d\omega
\end{aligned}
$$

Then, we combine J_2 with J_1 as

$$
\begin{aligned}
&J_1 + J_2 \\
&= 2\int_{\mathbb{S}^1} \nabla\Delta \log\frac{\rho_t}{e^{\omega\cdot\Omega_{\rho_t}}} \cdot \nabla \log\frac{\rho_t}{e^{\omega\cdot\Omega_{\rho_t}}}\rho_t \, d\omega \\
&+ \int_{\mathbb{S}^1} \nabla|\nabla \log\frac{\rho_t}{e^{\omega\cdot\Omega_{\rho_t}}}|^2 \cdot \nabla \log\frac{\rho_t}{e^{\omega\cdot\Omega_{\rho_t}}}\rho_t \, d\omega \\
&\quad + 2\int_{\mathbb{S}^1} \nabla\Big(\nabla(\omega\cdot\Omega_{\rho_t}) \cdot \nabla \log\frac{\rho_t}{e^{\omega\cdot\Omega_{\rho_t}}}\Big) \cdot \nabla \log\frac{\rho_t}{e^{\omega\cdot\Omega_{\rho_t}}}\rho_t \, d\omega \\
&= 2\int_{\mathbb{S}^1} \nabla\Delta \log\frac{\rho_t}{e^{\omega\cdot\Omega_{\rho_t}}} \cdot \nabla \log\frac{\rho_t}{e^{\omega\cdot\Omega_{\rho_t}}}\rho_t \, d\omega \\
&+ \int_{\mathbb{S}^1} \nabla|\nabla \log\frac{\rho_t}{e^{\omega\cdot\Omega_{\rho_t}}}|^2 \cdot \Big(\nabla\rho_t - \nabla(\omega\cdot\Omega_{\rho_t})\rho_t\Big) \, d\omega \\
&\quad + 2\int_{\mathbb{S}^1} \nabla\Big(\nabla(\omega\cdot\Omega_{\rho_t}) \cdot \nabla \log\frac{\rho_t}{e^{\omega\cdot\Omega_{\rho_t}}}\Big) \cdot \nabla \log\frac{\rho_t}{e^{\omega\cdot\Omega_{\rho_t}}}\rho_t \, d\omega \\
&= 2\int_{\mathbb{S}^1} \Big[\nabla\Delta \log\frac{\rho_t}{e^{\omega\cdot\Omega_{\rho_t}}} \cdot \nabla \log\frac{\rho_t}{e^{\omega\cdot\Omega_{\rho_t}}} - \frac{1}{2}\Delta|\nabla \log\frac{\rho_t}{e^{\omega\cdot\Omega_{\rho_t}}}|^2\Big]\rho_t \, d\omega \\
&\quad + 2\int_{\mathbb{S}^1} \Big[\nabla\Big(\nabla(\omega\cdot\Omega_{\rho_t}) \cdot \nabla \log\frac{\rho_t}{e^{\omega\cdot\Omega_{\rho_t}}}\Big) \cdot \nabla \log\frac{\rho_t}{e^{\omega\cdot\Omega_{\rho_t}}} \\
&- \frac{1}{2}\nabla|\nabla \log\frac{\rho_t}{e^{\omega\cdot\Omega_{\rho_t}}}|^2 \cdot \nabla(\omega\cdot\Omega_{\rho_t})\Big]\rho_t \, d\omega \\
&= -2\int \mathrm{tr}\Big(\Big[D^2 \log\frac{\rho_t}{e^{\omega\cdot\Omega_{\rho_t}}}\Big]^T D^2 \log\frac{\rho_t}{e^{\omega\cdot\Omega_{\rho_t}}}\Big)\rho_t \, d\omega \\
&\quad + 2\int \nabla \log\frac{\rho_t}{e^{\omega\cdot\Omega_{\rho_t}}} D^2(\omega\cdot\Omega_{\rho_t}) \nabla \log\frac{\rho_t}{e^{\omega\cdot\Omega_{\rho_t}}}\rho_t \, d\omega
\end{aligned}
$$

where we have used the Bochner formula on $\mathbb{S}^1$ to get the first integral in the last inequality.

On the other hand, by integration by parts, we get

$$
\begin{aligned}
J_3 &= 2\int_{\mathbb{S}^1} \Delta\rho_t \partial_t(\omega\cdot\Omega_{\rho_t})\,d\omega + 2\int_{\mathbb{S}^1} \nabla(\omega\cdot\Omega_{\rho_t})\cdot\nabla\rho_t\partial_t(\omega\cdot\Omega_{\rho_t})\,d\omega \\
&\quad + 2\int_{\mathbb{S}^1} \nabla(\omega\cdot\Omega_{\rho_t})\cdot\nabla\partial_t(\omega\cdot\Omega_{\rho_t})\rho_t\,d\omega - 2\int_{\mathbb{S}^1} \nabla(\omega\cdot\Omega_{\rho_t})\cdot\nabla\rho_t\partial_t(\omega\cdot\Omega_{\rho_t})\,d\omega \\
&= 2\int_{\mathbb{S}^1} \Delta\rho_t \partial_t(\omega\cdot\Omega_{\rho_t})\,d\omega + 2\int_{\mathbb{S}^1} \nabla(\omega\cdot\Omega_{\rho_t})\cdot\nabla\partial_t(\omega\cdot\Omega_{\rho_t})\rho_t\,d\omega \\
&= -2\int_{\mathbb{S}^1} \nabla\rho_t\cdot\nabla(\omega\cdot\partial_t\Omega_{\rho_t})\,d\omega + 2\int_{\mathbb{S}^1} \nabla(\omega\cdot\Omega_{\rho_t})\cdot\nabla(\omega\cdot\partial_t\Omega_{\rho_t})\rho_t\,d\omega \\
&= -2\int_{\mathbb{S}^1} \mathbb{P}_{\omega^\perp}\partial_t\Omega_{\rho_t}\cdot\nabla\log\frac{\rho_t}{e^{\omega\cdot\Omega_{\rho_t}}}\rho_t\,d\omega \\
&= -2\partial_t\Omega_{\rho_t}\cdot\int_{\mathbb{S}^1} \nabla\log\frac{\rho_t}{e^{\omega\cdot\Omega_{\rho_t}}}\rho_t\,d\omega,
\end{aligned}
$$

where we have used (C.2) and (C.3).

Then, we use (1.2) and (C.2)-(C.3) to compute $\partial_t\Omega_{\rho_t}$ as

$$
\begin{aligned}
\partial_t\Omega_{\rho_t} &= \frac{\partial_t J_{\rho_t}}{|J_{\rho_t}|} - \frac{J_{\rho_t}}{|J_{\rho_t}|^2}\frac{J_{\rho_t}\cdot\partial_t J_{\rho_t}}{|J_{\rho_t}|} = \frac{\partial_t J_{\rho_t}}{|J_{\rho_t}|} - \frac{\Omega_{\rho_t}}{|J_{\rho_t}|}\int_{\mathbb{S}^1}\omega\cdot\Omega_{\rho_t}\partial_t\rho_t\,d\omega \\
&= \frac{\partial_t J_{\rho_t}}{|J_{\rho_t}|} + \frac{\Omega_{\rho_t}}{|J_{\rho_t}|}\int_{\mathbb{S}^1}\nabla(\omega\cdot\Omega_{\rho_t})\cdot\nabla\log\frac{\rho_t}{e^{\omega\cdot\Omega_{\rho_t}}}\rho_t\,d\omega \qquad (3.6)\\
&= \frac{\partial_t J_{\rho_t}}{|J_{\rho_t}|} + \frac{\Omega_{\rho_t}}{|J_{\rho_t}|}\Omega_{\rho_t}\cdot\int_{\mathbb{S}^1}\nabla\log\frac{\rho_t}{e^{\omega\cdot\Omega_{\rho_t}}}\rho_t\,d\omega.
\end{aligned}
$$

Again, using (1.2) and (C.2)-(C.3), we have

$$
\begin{aligned}
J_3 &= -\frac{2}{|J_{\rho_t}|}\int_{\mathbb{S}^1}\Big(\omega\cdot\int_{\mathbb{S}^1}\nabla_{\omega'}\log\frac{\rho_t}{e^{\omega'\cdot\Omega_{\rho_t}}}\rho_t d\omega'\Big)\partial_t\rho_t\, d\omega\\
&\quad -\frac{2}{|J_{\rho_t}|}\Big(\Omega_{\rho_t}\cdot\int_{\mathbb{S}^1}\nabla\log\frac{\rho_t}{e^{\omega\cdot\Omega_{\rho_t}}}\rho_t\, d\omega\Big)^2\\
&= \frac{2}{|J_{\rho_t}|}\int_{\mathbb{S}^1}\nabla_\omega\Big(\omega\cdot\int_{\mathbb{S}^1}\nabla_{\omega'}\log\frac{\rho_t}{e^{\omega'\cdot\Omega_{\rho_t}}}\rho_t d\omega'\Big)\cdot\nabla_\omega\log\frac{\rho_t}{e^{\omega\cdot\Omega_{\rho_t}}}\rho_t d\omega\\
&\quad -\frac{2}{|J_{\rho_t}|}\Big(\Omega_{\rho_t}\cdot\int_{\mathbb{S}^1}\nabla\log\frac{\rho_t}{e^{\omega\cdot\Omega_{\rho_t}}}\rho_t\, d\omega\Big)^2\\
&= \frac{2}{|J_{\rho_t}|}\int_{\mathbb{S}^1}\mathbb{P}_{\omega^\perp}\int_{\mathbb{S}^1}\nabla_{\omega'}\log\frac{\rho_t}{e^{\omega'\cdot\Omega_{\rho_t}}}\rho_t d\omega'\cdot\nabla_\omega\log\frac{\rho_t}{e^{\omega\cdot\Omega_{\rho_t}}}\rho_t d\omega\\
&\quad -\frac{2}{|J_{\rho_t}|}\Big(\Omega_{\rho_t}\cdot\int_{\mathbb{S}^1}\nabla\log\frac{\rho_t}{e^{\omega\cdot\Omega_{\rho_t}}}\rho_t\, d\omega\Big)^2\\
&= \frac{2}{|J_{\rho_t}|}\bigg(\Big|\int_{\mathbb{S}^1}\nabla\log\frac{\rho_t}{e^{\omega\cdot\Omega_{\rho_t}}}\rho_t d\omega\Big|^2-\Big(\int_{\mathbb{S}^1}\Omega_{\rho_t}\cdot\nabla\log\frac{\rho_t}{e^{\omega\cdot\Omega_{\rho_t}}}\rho_t d\omega\Big)^2\bigg),
\end{aligned}
$$

where in the second equality, we have used (C.2) and (C.3).

Hence we complete (3.5). □

3.1 Proof of (3.3)

For the notational simplification, for any fixed constant ε_* satisfying $0<\varepsilon_*\le\frac{1}{10}$,

$$
C_* := \frac{2\pi^2 e^{2(1+|\log C_M|)}(1+\frac{1}{15}\varepsilon_*)}{1-\frac{7}{6}\varepsilon_*},
$$

denotes the coefficient of the logarithmic Sobolev inequality in Lemma 2.1 when $\lambda = 1+|\log C_M|$ because of $\Psi=-\omega\cdot\Omega-\log C_M$ in our case. Moreover, we set

$$
T_* := 2H(\rho_0|M_{\Omega_{\rho_0}})\Big[\min\Big(C_M^2e^{-2}\varepsilon_*^2, \frac{L}{2C_*}C_M^2e^{-2}\varepsilon_*^2, mC_*^{-1}\Big)\Big]^{-1}, \tag{3.7}
$$

where $L := (2+\frac{4}{m})^{-1}\log 2$, and $m := \left|\int_{\mathbb{S}^1} \omega M_{\Omega_\rho}\, d\omega\right|$.

Since $H(\rho_t|M_{\Omega_{\rho_t}}) \geq 0$ for all t by Jensen's inequality, (3.1) yields that for any $t > 0$,

$$\begin{aligned} H(\rho_0|M_{\Omega_{\rho_0}}) &\geq H(\rho_0|M_{\Omega_{\rho_0}}) - H(\rho_t|M_{\Omega_{\rho_t}}) = -\int_0^t \frac{d}{ds} H(\rho_s|M_{\Omega_{\rho_s}}) ds \\ &= \int_0^t I(\rho_s|M_{\Omega_{\rho_s}}) ds \geq t \min_{s\in[0,t]} I(\rho_s|M_{\Omega_{\rho_s}}). \end{aligned} \tag{3.8}$$

(We remark that Lemma 3.1 implies that the dissipation is lower semicontinuous as a function of time and consequently it achieves a minimum on any finite closed interval.)

Thus, there exists $T_0 \in [0, T_*]$ such that

$$I(\rho_{T_0}|M_{\Omega_{\rho_{T_0}}}) \leq \min_{s\in[0,T_*]} I(\rho_s|M_{\Omega_{\rho_s}}) \leq \frac{H(\rho_0|M_{\Omega_{\rho_0}})}{T_*} \leq \frac{1}{2} C_M^2 e^{-2} \varepsilon_*^2. \tag{3.9}$$

Therefore, it follows from (3.2) that

$$\|\rho_{T_0} M^{-1}_{\Omega_{\rho_{T_0}}} - 1\|_{L^\infty(\mathbb{S}^1)} \leq \varepsilon_*,$$

which together with Lemma 2.1 implies that

$$H(\rho_{T_0}|M_{\Omega_{\rho_{T_0}}}) \leq C_* I(\rho_{T_0}|M_{\Omega_{\rho_{T_0}}}).$$

Then, it follows from (3.7) and (3.9) that

$$H(\rho_{T_0}|M_{\Omega_{\rho_{T_0}}}) \leq C_* \frac{H(\rho_0|M_{\Omega_{\rho_0}})}{T_*} \leq \frac{m}{2}. \tag{3.10}$$

In particular, for all $t \geq T_0$,

$$H(\rho_t|M_{\Omega_{\rho_t}}) \leq \frac{m}{2}.$$

Then, using Csiszar-Kullback-Pinsker inequality (see for example [51, Theorem 1.4]), we have that for all $t \geq T_0$,

$$\begin{aligned} |J_{\rho_t}| &\geq \left| \int_{\mathbb{S}^1} \omega \cdot M_{\Omega_{\rho_t}} \, d\omega \right| - \left| J_{\rho_t} - \int_{\mathbb{S}^1} \omega \cdot M_{\Omega_{\rho_t}} \, d\omega \right| \\ &= m - \left| \int_{\mathbb{S}^1} \omega \left(\rho_t - M_{\Omega_{\rho_t}} \right) d\omega \right| \geq m - \| \rho_t - M_{\Omega_{\rho_t}} \|_{L^1(\mathbb{S}^1)} \geq \frac{m}{2}, \end{aligned} \tag{3.11}$$

which together with Lemma 3.1 implies that for any $t \geq s \geq T_0$,

$$I(\rho_t | M_{\Omega_{\rho_t}}) \leq I(\rho_s | M_{\Omega_{\rho_s}}) e^{(2+\frac{4}{m})(t-s)}. \tag{3.12}$$

Therefore, using (3.9), we have that for any $t \in [T_0, T_0 + L]$ (recall $L = (2 + \frac{4}{m})^{-1} \log 2$),

$$I(\rho_t | M_{\Omega_{\rho_t}}) \leq I(\rho_{T_0} | M_{\Omega_{\rho_{T_0}}}) e^{(2+\frac{4}{m})L} = 2 I(\rho_{T_0} | M_{\Omega_{\rho_{T_0}}}) \leq C_M^2 e^{-2} \varepsilon_*^2. \tag{3.13}$$

Now, we extend the time length, on which the above inequality still holds. Using the same estimates as (3.8), we have

$$\begin{aligned} H(\rho_{T_0} | M_{\Omega_{\rho_{T_0}}}) \geq H(\rho_{T_0+L/2} | M_{\Omega_{\rho_{T_0+L/2}}}) &\geq \int_{T_0+L/2}^{T_0+L} I(\rho_s | M_{\Omega_{\rho_s}}) ds \\ &\geq \frac{L}{2} \min_{s \in [T_0+L/2, T_0+L]} I(\rho_s | M_{\Omega_{\rho_s}}). \end{aligned}$$

Then, thanks to (3.7) and (3.10), there exists $s_* \in [T_0 + L/2, T_0 + L]$ such that

$$I(\rho_{s_*} | M_{\Omega_{\rho_{s_*}}}) \leq \min_{s \in [T_0+L/2, T_0+L]} I(\rho_s | M_{\Omega_{\rho_s}}) \leq \frac{2}{L} \frac{C_* H(\rho_0 | M_{\Omega_{\rho_0}})}{T_*} \leq \frac{1}{2\pi} C_M^2 e^{-2} \varepsilon_*^2.$$

Thus, the inequality (3.12) yields that for any $t \in [s_*, s_* + L]$,

$$I(\rho_t | M_{\Omega_{\rho_t}}) \leq I(\rho_{s_*} | M_{\Omega_{\rho_{s_*}}}) e^{(2+\frac{4}{m})L} = 2 I(\rho_{s_*} | M_{\Omega_{\rho_{s_*}}}) \leq C_M^2 e^{-2} \varepsilon_*^2,$$

which together with (3.13) implies that

$$I(\rho_t|M_{\Omega_{\rho_t}}) \le C_M^2 e^{-2}\varepsilon_*^2, \quad \forall t \in [T_0, T_0 + \frac{3}{2}L].$$

We use again the above arguments to make the above inequality hold on the interval $[T_0, T_0 + 2L]$, therefore we repeat the same arguments to eventually complete (3.3).

3.2 Proof of (1.7)

First of all, it follows from (3.2) and (3.3) that

$$\|\rho M_{\Omega_\rho}^{-1} - 1\|_{L^\infty(\mathbb{S}^1)} \le \varepsilon_*, \quad \forall t \ge T_0.$$

Then, Lemma 2.1 and (3.1) imply that

$$H(\rho_t|M_{\Omega_{\rho_t}}) \le C_* I(\rho_t|M_{\Omega_{\rho_t}}), \quad \forall t \ge T_0. \tag{3.14}$$

Now, using (1.9) and Lemma 3.1, we have that for any $0 < t < T_0$,

$$\begin{aligned} H(\rho_t|M_{\Omega_{\rho_t}}) - H(\rho_{T_0}|M_{\Omega_{\rho_{T_0}}}) &= -\int_t^{T_0} \frac{d}{ds} H(\rho_s|M_{\Omega_{\rho_s}})ds = \int_t^{T_0} I(\rho_s|M_{\Omega_{\rho_s}})ds \\ &\le I(\rho_t|M_{\Omega_{\rho_t}}) \int_t^{T_0} \exp\left[\left(2 + \frac{2}{|J_{\rho_0}|e^{-T_0}}\right)s\right]ds \\ &= I(\rho_t|M_{\Omega_{\rho_t}}) \\ &\times \underbrace{\frac{|J_{\rho_0}|e^{-T_0}}{2|J_{\rho_0}|e^{-T_0} + 2}\left(\exp\left[\left(2 + \frac{2}{|J_{\rho_0}|e^{-T_0}}\right)T_0\right] - \exp\left[\left(2 + \frac{2}{|J_{\rho_0}|e^{-T_0}}\right)t\right]\right)}_{=:\beta(t)}. \end{aligned}$$

Thus, (3.14) and Lemma 3.1 yield that for any $0 < t < T_0$,

$$\begin{aligned} H(\rho_t|M_{\Omega_{\rho_t}}) &\le I(\rho_t|M_{\Omega_{\rho_t}})\beta(t) + H(\rho_{T_0}|M_{\Omega_{\rho_{T_0}}}) \\ &\le I(\rho_t|M_{\Omega_{\rho_t}})\beta(t) + C_* I(\rho_{T_0}|M_{\Omega_{\rho_{T_0}}}) \\ &\le \left(\beta(t) + C_* \exp\left[\left(2 + \frac{2}{|J_{\rho_0}|e^{-T_0}}\right)(T_0 - t)\right]\right) I(\rho_t|M_{\Omega_{\rho_t}}), \end{aligned}$$

which together with (3.14) implies that

$$H(\rho_t|M_{\Omega_{\rho_t}}) \le \frac{1}{B(t)} I(\rho_t|M_{\Omega_{\rho_t}}), \quad \forall t > 0,$$

where

$$\begin{aligned} B(t) := \Bigg(\frac{|J_{\rho_0}|e^{-T_0}}{2|J_{\rho_0}|e^{-T_0} + 2} \Bigg(&\exp\left[\left(2 + \frac{2}{|J_{\rho_0}|e^{-T_0}}\right) T_0\right] \\ &- \exp\left[\left(2 + \frac{2}{|J_{\rho_0}|e^{-T_0}}\right) t\right]\Bigg)_+ \\ &+ C_* \exp\left[\left(2 + \frac{2}{|J_{\rho_0}|e^{-T_0}}\right)(T_0 - t)_+\right]\Bigg)^{-1}. \end{aligned} \tag{3.15}$$

Therefore, it follows from (3.1) that

$$\frac{d}{dt} H(\rho_t|M_{\Omega_{\rho_t}}) \le -B(t) H(\rho_t|M_{\Omega_{\rho_t}}),$$

which implies that

$$H(\rho_t|M_{\Omega_{\rho_t}}) \le H(\rho_0|M_{\Omega_{\rho_0}}) e^{-\int_0^t B(s)ds}.$$

Then, the Csiszar-Kullback-Pinsker inequality yields that

$$\|\rho_t - M_{\Omega_{\rho_t}}\|_{L^1(\mathbb{S}^1)} \le H(\rho_0|M_{\Omega_{\rho_0}}) e^{-\int_0^t B(s)ds}, \quad \forall t \ge 0. \tag{3.16}$$

Following the same estimates as in proof of [44, Theorem 2.2], there exists a constant $C > 0$ such that

$$\left|\frac{d}{dt}J_{\rho_t}\right| \le C\|\rho_t - M_{\Omega_{\rho_t}}\|_{L^1(\mathbb{S}^1)},$$

therefore,

$$\left|\frac{d}{dt}J_{\rho_t}\right| \le CH(\rho_0|M_{\Omega_{\rho_0}})e^{-\int_0^t B(s)ds}.$$

Notice that since $e^{-\int_0^t B(s)ds} \to 0$ as $t \to \infty$, and thanks to (1.9), there exists a constant vector J_∞ such that $|J_\infty| \ge \frac{m}{2}$ and

$$\begin{aligned}|J_{\rho_t} - J_\infty| &\le \int_t^\infty \left|\frac{d}{dr}J_{\rho_r}\right| dr \\ &\le CH(\rho_0|M_{\Omega_{\rho_0}}) \int_t^\infty e^{-\int_0^r B(s)ds} dr.\end{aligned}$$

Then, setting $\Omega_\infty = \frac{J_\infty}{|J_\infty|}$, we have

$$\begin{aligned}|\Omega_{\rho_t} - \Omega_\infty| &\le \frac{2|J_{\rho_t} - J_\infty|}{|J_\infty|} \le \frac{4}{m}|J_{\rho_t} - J_\infty| \\ &\le C\frac{4}{m}H(\rho_0|M_{\Omega_{\rho_0}}) \int_t^\infty e^{-\int_0^r B(s)ds} dr.\end{aligned}$$

Let $\gamma_t : [0,1] \to \mathbb{R}^2$ be a curve defined by

$$\gamma_t(s) = (1-s)\Omega_{\rho_t} + s\Omega_\infty.$$

Then, we have

$$\begin{aligned}\|M_{\Omega_{\rho_t}} - M_{\Omega_\infty}\|_{L^1(\mathbb{S}^1)} &= C_M \int_{\mathbb{S}^1} |e^{\omega\cdot\Omega_{\rho_t}} - e^{\omega\cdot\Omega_\infty}| d\omega \\ &\le C_M \int_{\mathbb{S}^1}\int_0^1 \left|\frac{d}{ds}e^{\omega\cdot\gamma_t(s)}\right| ds\, d\omega \\ &\le C_M|\Omega_{\rho_t} - \Omega_\infty| \int_{\mathbb{S}^1}\int_0^1 e^{\omega\cdot\gamma_t(s)} ds d\omega \le 2\pi e^2 C_M |\Omega_{\rho_t} - \Omega_\infty| \\ &\le \frac{8\pi^2 e^2 CC_M}{m} H(\rho_0|M_{\Omega_{\rho_0}}) \int_t^\infty e^{-\int_0^r B(s)ds} dr.\end{aligned}$$

Hence, we combine the above estimate with (3.16), to get

$$\begin{aligned}\|\rho_t - M_{\Omega_\infty}\|_{L^1(\mathbb{S}^1)} &\le \|\rho_t - M_{\Omega_{\rho_t}}\|_{L^1(\mathbb{S}^1)} + \|M_{\Omega_{\rho_t}} - M_{\Omega_\infty}\|_{L^1(\mathbb{S}^1)} \\ &\le H(\rho_0|M_{\Omega_{\rho_0}})\left(e^{-\int_0^t B(s)ds} + \frac{8\pi^2 e^2 C C_M}{m}\int_t^\infty e^{-\int_0^r B(s)ds} dr\right).\end{aligned} \tag{3.17}$$

3.3 Conclusion

We now use (3.17) to show the three results 1)-3) in Theorem 1.1. Since the function $B(t)$ is positive and non-decreasing, (3.17) implies that

$$\begin{aligned}\|\rho_t - M_{\Omega_\infty}\|_{L^1(\mathbb{S}^1)} &\le H(\rho_0|M_{\Omega_{\rho_0}})\left(e^{-B(0)t} + C\int_t^\infty e^{-B(0)r} dr\right) \\ &\le CH(\rho_0|M_{\Omega_{\rho_0}})e^{-B(0)t},\end{aligned}$$

where it follows from (3.15) that

$$\begin{aligned}B(0) = \Bigg(\frac{|J_{\rho_0}|e^{-T_0}}{2|J_{\rho_0}|e^{-T_0} + 2}\left(\exp\left[\left(2 + \frac{2}{|J_{\rho_0}|e^{-T_0}}\right)T_0\right] - 1\right) \\ + C_* \exp\left[\left(2 + \frac{2}{|J_{\rho_0}|e^{-T_0}}\right)T_0\right]\Bigg)^{-1}.\end{aligned}$$

On the other hand, using (3.17) together with the fact that $B(t) = C_*^{-1}$ for all $t \ge T_0$, we have that

$$\|\rho_t - M_{\Omega_\infty}\|_{L^1(\mathbb{S}^1)} \le CH(\rho_0|M_{\Omega_{\rho_0}})e^{-C_*^{-1}(t-T_0)}, \quad \forall t \ge T_0.$$

In order to show the last result, we observe that $C_*(\varepsilon_*) \to 2\pi^2 e^{2(1+|\log C_M|)}$ as $\varepsilon_* \to 0$. Thus, for any $\varepsilon > 0$, there exists $\tilde{\varepsilon}_*$ such that

$$C_*^{-1}(\tilde{\varepsilon}_*) \ge \frac{1}{2\pi^2 e^{2(1+|\log C_M|)}} - \frac{\varepsilon}{2}.$$

For such $\tilde{\varepsilon}_*$, since

$$T_0 \le 2H(\rho_0|M_{\Omega_{\rho_0}})\left[\min\left(C_M^2 e^{-2}\tilde{\varepsilon}_*^2, \frac{L}{2C_*(\tilde{\varepsilon}_*)}C_M^2 e^{-2}\tilde{\varepsilon}_*^2, mC_*^{-1}(\tilde{\varepsilon}_*)\right)\right]^{-1},$$

if $H(\rho_0|M_{\Omega_{\rho_0}}) \to 0$, then $T_0 \to 0$ and $|J_{\rho_0}| \to m$, consequently

$B(0) \to C_*^{-1}(\tilde{\varepsilon}_*)$.

Thus, there exists $\delta > 0$ such that if $H(\rho_0|M_{\Omega_{\rho_0}}) < \delta$, then

$$B(0) \geq C_*^{-1}(\tilde{\varepsilon}_*) - \frac{\varepsilon}{2}.$$

Therefore, we have

$$\begin{aligned}\|\rho_t - M_{\Omega_\infty}\|_{L^1(\mathbb{S}^1)} &\leq CH(\rho_0|M_{\Omega_{\rho_0}})e^{-B(0)t}\\ &\leq CH(\rho_0|M_{\Omega_{\rho_0}}) \exp\Big[-\Big(\frac{1}{2\pi^2 e^{2(1+|\log C_M|)}} - \varepsilon\Big)t\Big].\end{aligned}$$

Hence we complete the proof.

Appendix

A Laplacian log identity

Lemma A.1. *Assume ρ and Ψ are functions in $\mathcal{C}^2(\mathbb{S}^1)$, and $\rho > 0$. Then, the following identity holds:*

$$\frac{\Delta\rho}{\rho} + \frac{\nabla\Psi \cdot \nabla\rho}{\rho} + \Delta\Psi = \Delta \log \frac{\rho}{e^{-\Psi}} + \Big|\nabla \log \frac{\rho}{e^{-\Psi}}\Big|^2 - \nabla\Psi \cdot \nabla \log \frac{\rho}{e^{-\Psi}}, \quad \text{(A.1)}$$

Proof. We begin by computing the first term in r.h.s of (A.1) as

$$\Delta \log \frac{\rho}{e^{-\Psi}} = \mathrm{div} \frac{\nabla \frac{\rho}{e^{-\Psi}}}{\frac{\rho}{e^{-\Psi}}} = -\frac{\Big|\nabla \frac{\rho}{e^{-\Psi}}\Big|^2}{\Big(\frac{\rho}{e^{-\Psi}}\Big)^2} + \frac{\Delta \frac{\rho}{e^{-\Psi}}}{\frac{\rho}{e^{-\Psi}}} =: I_1 + I_2.$$

By straightforward computations, we have

$$I_1 = -\Big|\nabla \log \frac{\rho}{e^{-\Psi}}\Big|^2,$$

and

$$
\begin{aligned}
I_2 &= \frac{\operatorname{div}\left(\frac{\nabla\rho}{e^{-\Psi}} + \frac{\rho}{e^{-\Psi}}\nabla\Psi\right)}{\frac{\rho}{e^{-\Psi}}} \\
&= \frac{\Delta\rho}{\rho} + 2\frac{\nabla\Psi\cdot\nabla\rho}{\rho} + \Delta\Psi + |\nabla\Psi|^2 \\
&= \frac{\Delta\rho}{\rho} + \frac{\nabla\Psi\cdot\nabla\rho}{\rho} + \Delta\Psi + \nabla\Psi\cdot\nabla\log\frac{\rho}{e^{-\Psi}}.
\end{aligned}
$$

Hence, we have the desired identity. □

B Uniqueness

We here present another proof for the uniqueness of weak solutions to (1.1), combining the short-time stability (1.10) with the energy estimates to be specified below. First of all, we see that the weak solution ρ to (1.1) becomes smooth instantaneously thanks to the parabolic regularity, therefore we have that $\rho \in L^\infty([t_0,\infty)\times\mathbb{S}^1)$ for any fixed $t_0 > 0$.

Let ρ and $\bar\rho$ be any two weak solutions to (1.1). Then, since

$$
\partial_t(\rho-\bar\rho) = \Delta_\omega(\rho-\bar\rho) - \nabla_\omega\cdot\Big((\rho-\bar\rho)\,\mathbb{P}_{\omega^\perp}\Omega_\rho\Big) - \nabla_\omega\cdot\Big(\bar\rho\,\mathbb{P}_{\omega^\perp}(\Omega_\rho-\Omega_{\bar\rho})\Big),
$$

we have

$$
\begin{aligned}
\frac{d}{dt}\int_{\mathbb{S}^1}\frac{|\rho-\bar\rho|^2}{2}d\omega + \int_{\mathbb{S}^1}|\nabla_\omega(\rho-\bar\rho)|^2 d\omega &= \int_{\mathbb{S}^1}(\rho-\bar\rho)\Omega_\rho\cdot\nabla_\omega(\rho-\bar\rho)d\omega \\
&\quad + \int_{\mathbb{S}^1}\bar\rho(\Omega_\rho-\Omega_{\bar\rho})\cdot\nabla_\omega(\rho-\bar\rho)d\omega.
\end{aligned}
$$

Since (1.9) yields that

$$
|\Omega_\rho-\Omega_{\bar\rho}| \le \frac{2|J_\rho-J_{\bar\rho}|}{|J_\rho|} \le Ce^t|J_{\rho_0}|^{-1}\|\rho-\bar\rho\|_{L^2(\mathbb{S}^1)},
$$

we have

$$\left| \int_{\mathbb{S}^1} \bar{\rho}(\Omega_\rho - \Omega_{\bar{\rho}}) \cdot \nabla_\omega (\rho - \bar{\rho}) d\omega \right| \leq Ce^t |J_{\rho_0}|^{-1} \|\bar{\rho}\|_{L^\infty(\mathbb{S}^1)} \|\rho - \bar{\rho}\|_{L^2(\mathbb{S}^1)} \|\nabla_\omega(\rho - \bar{\rho})\|_{L^2(\mathbb{S}^1)}.$$

Then, using $\rho \in L^\infty([t_0, \infty) \times \mathbb{S}^1)$ for all $t \geq t_0$, we have that for all $t \geq t_0$

$$\frac{d}{dt} \int_{\mathbb{S}^1} |\rho - \bar{\rho}|^2 d\omega + \int_{\mathbb{S}^1} |\nabla_\omega (\rho - \bar{\rho})|^2 d\omega \leq C(1 + e^{2t}) \int_{\mathbb{S}^1} |\rho - \bar{\rho}|^2 d\omega,$$

which implies that

$$\int_{\mathbb{S}^1} |\rho - \bar{\rho}|^2 d\omega \leq Ce^{e^{2t}} \int_{\mathbb{S}^1} |\rho_{t_0} - \bar{\rho}_{t_0}|^2 d\omega, \quad \forall t \geq t_0.$$

Therefore, if $\rho_{t_0} = \bar{\rho}_{t_0}$, then $\rho_t = \bar{\rho}_t$ for all $t \geq t_0$.
Since it follows from the short-time stability (1.10) that if $\rho_0 = \bar{\rho}_0$, then $\rho_t = \bar{\rho}_t$ for all $t \leq \delta$, we take $t_0 < \delta$ to complete the uniqueness.

C Formulas for Calculus on sphere

We here present some useful formulas on n-dimensional sphere $\mathbb{S}^n$, which are extensively used in this paper.
Let F be a vector-valued function and f be scalar-valued function. Then we have the following formulas related to the integration by parts:

$$\int_{\mathbb{S}^n} f \nabla_\omega \cdot F d\omega = -\int_{\mathbb{S}^n} F \cdot (\nabla_\omega f - 2\omega f) d\omega. \tag{C.1}$$

By the definition of the projection $\mathbb{P}_{\omega^\perp}$, it is obvious that

$$\mathbb{P}_{\omega^\perp} \nabla_\omega f = \nabla_\omega f \tag{C.2}$$

for any scalar-valued function f.

Moreover, for any constant vector $v \in \mathbb{R}^{n+1}$, we have

$$\nabla_\omega(\omega \cdot v) = \mathbb{P}_{\omega^\perp} v. \tag{C.3}$$

We refer to [47, 85] for the derivations of these formulas.

Chapter 5

Emergence of phase concentration for the Kuramoto-Sakaguchi equation

We study the asymptotic phase concentration phenomena for the Kuramoto-Sakaguchi(K-S) equation in a large coupling strength regime. For this, we analyze the detailed dynamics of the order parameters such as the amplitude and the average phase. For the infinite ensemble of oscillators with the identical natural frequency, we show that the total mass distribution concentrates on the average phase asymptotically, whereas the mass around the antipodal point of the average phase decays to zero exponentially fast in any positive coupling strength regime. Thus, generic initial kinetic densities evolve toward the Dirac measure concentrated on the average phase. In contrast, for the infinite ensemble with distributed natural frequencies, we find a certain time-dependent interval whose length can be explicitly quantified in terms of the coupling strength. Provided that the coupling strength is sufficiently large, the mass on such an interval is eventually non-decreasing over the time. We also show that the amplitude order parameter has a positive lower bound that depends on the size of support of the distribution function for the natural frequencies and the coupling strength. The proposed asymptotic lower bound on the order parameter tends to unity, as the coupling strength increases to infinity. This is reminiscent of practical synchronization for the Kuramoto model, in which the diameter for the phase configuration is inversely proportional to the coupling strength. Our results for the K-S equation generalize the results in [57] on the emergence of phase-locked states for the Kuramoto model in a large coupling strength regime.

1 Introduction

Collective phenomena such as aggregation, flocking, and synchronization, etc., are ubiquitous in biological, chemical, and mechanical systems in nature, e.g., the flashing of fireflies, chorusing of crickets, synchronous firing of cardiac pacemakers, and metabolic synchrony in yeast cell suspensions (see for instance [1, 18]). After Huygens' observation on the anti-synchronized motion of two pendulum clocks hanging on the same bar, the synchronization of oscillators were reported in literature from time to time. However, the first rigorous and systematic studies on synchronization were pioneered by Winfree [95] and Kuramoto [70] in several decades ago. They introduced phase coupled models for the ensemble of weakly coupled oscillators, and showed that collective synchronization in the ensemble of oscillators can emerge from disordered ensemble via the competing mechanism between intrinsic randomness and sinusoidal nonlinear couplings (see [1, 36, 87], for details). In this paper, we are interested in the large-time dynamics of a large ensemble of Kuramoto oscillators. In particular, we assume that the number of Kuramoto oscillators is sufficiently large so that a one-oscillator probability distribution function can describe effectively the dynamics of a large phase-coupled system, i.e., our main concern lies in the mesoscopic description of the ensemble of Kuramoto oscillators. In fact, this kinetic description has been used in physics literature [1] to analyze the phase transition from an incoherent state to a partially synchronized state, as the coupling strength is varied from zero to a large value.

Let $f = f(\theta, \omega, t)$ be the one-oscillator probability density function of the Kuramoto ensemble in phase $\theta \in \mathbb{T} := \mathbb{R}/(2\pi\mathbb{Z})$, with a natural frequency ω at time t, as in [71]. Suppose that $g = g(\omega)$ is a nonnegative and compactly supported probability density function for natural frequencies with zero first frequency moment ($\int_{\mathbb{R}} \omega g(\omega) d\omega = 0$). Then, the dynamics of the kinetic density f is governed by the Kuramoto-Sakaguchi (K-S) equation:

$$\begin{cases} \partial_t f + \partial_\theta (v[f] f) = 0, \qquad (\theta, \omega) \in \mathbb{T} \times \mathbb{R}, \quad t > 0, \\ v[f](\theta, \omega, t) = \omega - K \displaystyle\int_{\mathbb{T}} \sin(\theta - \theta_*) \rho(\theta_*, t) \, d\theta_*, \quad \rho(\theta, t) := \displaystyle\int_{\mathbb{R}} f(\theta, \omega, t) \, d\omega, \end{cases} \tag{1.1}$$

subject to the initial datum:

$$f(\theta, \omega, 0) = f_0(\theta, \omega), \quad \int_{\mathbb{T}} f_0 \, d\theta = g(\omega), \tag{1.2}$$

where K is the positive coupling strength measuring the degree of mean-field interactions between oscillators. The K-S equation (1.1) has been rigorously derived from the Kuramoto model in mean-field limit ($N \to \infty$), using the method of particle-in-cell employing empirical measures as an approximation [71]. Several global existence theories have been proposed for (1.1)-(1.2) in different frameworks, e.g., BV-entropic weak solutions [2], measure-valued solutions, and classical solutions [20, 22, 71]. Recently, motivated by the success of nonlinear Landau damping in plasma physics, there have been several interesting works [14, 23, 41, 96] on the Kuramoto conjecture and Landau damping in relation to stability and instability of incoherent solution in sub-and supercritical regimes. We also refer to [34, 61, 62] for the corresponding issues for the Kuramoto-Sakaguchi-Fokker-Planck equation which is a stochastic version of the K-S equation.

The purpose of this paper is to investigate the emergence of phase concentration for the K-S equation via the time-asymptotic approach. The time-asymptotic approach is to show existence of the steady states with some desired properties, as well as their stability to a given time-dependent problem. This approach has been very successful in the realms of hyperbolic conservation laws and kinetic theory, to analyze the large-time behavior of viscous conservation laws and positivity of Boltzmann shocks [73, 74]. In spirit, this is close to the mean curvature flow in differential geometry, in which manifolds with constant mean curvature emerges as an asymptotic manifold from a rough manifold via the mean curvature flow. In this time-asymptotic approach, we are able to obtain quantitive estimates on the detailed relaxation dynamics from the initial states not in the resulting attractors. As byproduct, stability and structure of the resulting attractors follows; for a finite-dimensional analogue, we refer to [20] where existence, stability and structure of the phase-locked states are presented via the time-asymptotic approach based on the Kuramoto model. For a survey on related issues arising from the classical and quantum synchronization, we refer to the recent review papers [36, 58].

The main results of this paper are three-fold. First, we consider the infinite ensemble of identical oscillators in which the density function $g = g(\omega)$

for the natural frequency is given by the Dirac measure concentrated on the average natural frequency. In this case, for any positive coupling strength, we show that generic $\mathcal{C}^1$- initial datum with a positive order parameter tends to the Dirac measure concentrated on the asymptotic average phase, whereas mass near the antipodal phase of the average phase decays to zero exponentially fast (see Theorem 3.1). The latter assertion contrasts the difference between the infinite-dimensional case (the K-S equation) and finite-dimensional case (the Kuramoto model). For the Kuramoto model, bi-polar configurations (say, one oscillator lies on the south pole, and the rest of ensemble lies on the north pole) is possible, although it is unstable. The second and third results deal with mass concentration phenomenon for the distributed natural frequencies. In our second result (Theorem 3.2), we construct a time dependent interval $L(t)$ centered at the time-dependent average phase and with constant width, such that the mass over $L(t)$ is nondecreasing and for each fixed natural frequency w in the support of g, the integral $\int_{L(t)} |f(\theta,\omega,t)|^2 d\theta$ tends to infinity exponentially fast for large coupling strengths depending on the size of the support of g. This is obtained for a well arranged initial datum. Such condition is removed in our third result, where we present a nontrivial lower bound for the asymptotic amplitude order parameter depending only on the size of the support of g and the coupling strength. We also show that there exists a time dependent interval that contains all the mass asymptotically. The size of the interval is characterized by the coupling strength, maximum of natural frequencies, and the asymptotic amplitude order parameter (see Theorem

3.3). Moreover, by choosing the coupling strength large enough, this size can be made arbitrary small, and the amplitude order parameter arbitrary close to 1.

The rest of this paper is organized as follows. In Section 2, we briefly review several concepts of synchronization for the Kuramoto model, the order parameters (amplitude and phase), and gradient flow formulations for the Kuramoto model and the K-S equation. We also recall some relevant previous results for the Kuramoto model. In Section 3, we discuss our main results for the K-S equation on the emergence of attractors. In Section 4, we present an emergent dynamics of the K-S equation for identical oscillators. In particular, we present dynamics of the amplitude order parameter and using it, we give the proof of Theorem 3.1. In Section 5, we study the dynamics of local order parameters for the sub-ensemble of identical oscillators with the same natural frequency, and using the detailed dynamics of local order parameters, we provide the proof of Theorem 3.2. In Section 6, we provide a nontrivial lower bound for the asymptotic amplitude order parameter in terms of the maximum of the natural frequency, and the coupling strength. This lower bound estimate for the asymptotic order parameter yields a certain practical synchronization that has been introduced for the finite-dimensional Kuramoto model in [60]. Finally, Section 7 is devoted to the summary of our main results and future directions. In Appendix A, we provide a short presentation of Otto's calculus which inspired us for some of our proofs. For the readability of the paper, we postpone lengthy proofs of several lemmata and propositions to Appendix B - Appendix E. In Appendix F, we discuss several estimates on the Kuramoto vector field.

Notation: For vectors p, q in $\mathbb{R}^2$, we denote an inner product of p and q by $p \cdot q$, whereas for two complex numbers $z_1, z_2 \in \mathbb{C}$, we set their inner product by $\langle z_1, z_2 \rangle = z_1 \bar{z}_2$.

2 Preliminaries

In this section, we briefly review two synchronization models, the Kuramoto model and its kinetic counterpart, the Kuramoto-Sakaguchi equation. For these two models, we introduce real-valued order parameters and gradient flow formulations.

2.1 The Kuramoto model

Consider a complete network consisting of N-nodes with edges connecting all pair of nodes, and assume that at each node, a Landau-Stuart oscillator is located. We set $z_i \in \mathbb{C}^1$ to be the state of the i-th Landau-Stuart oscillator. Then, z_i is governed by the following first-order system of ODEs:

$$\frac{dz_i}{dt} = (1 - |z_i|^2 + \mathrm{i}\omega_i) z_i + \frac{K}{N} \sum_{j=1}^{N} (z_j - z_i), \quad j = 1, \cdots, N, \tag{2.1}$$

where K is the uniform coupling strength between oscillators, and ω_i is the quenched random natural frequency of the i-th Stuart-Landau oscillator extracted from a given distribution function $g = g(\omega)$, $\omega \in \mathbb{R}$:

$$\int_{\mathbb{R}} g(\omega)\; d\omega = 1, \quad \int_{\mathbb{R}} \omega g(\omega)\; d\omega = 0, \quad \text{supp } g(\cdot) \subset\subset \mathbb{R}, \quad g(\omega) \geq 0.$$

The state $z_i = z_i(t)$ governed by the system (2.1) approaches a certain limit-cycle (a circle with radius determined by the coupling strength) asymptotically for a suitable range of K (see [69]). Hence, in the sequel, we are mainly interested in the dynamics of the limit-cycle oscillators so that the amplitude variations can be ignored from the dynamics, and we focus our attention on the phase dynamics. This explains the meaning of "*weakly coupled* oscillator". To see the dynamics of the phase, we set

$$z_i(t) := e^{\mathrm{i}\theta_i(t)}, \quad t \geq 0, \quad 1 \leq i \leq N, \tag{2.2}$$

and substitute (2.2) into (2.1), and compare the imaginary part of the resulting relation to derive the Kuramoto model [69, 70]:

$$\dot{\theta}_i = \omega_i + \frac{K}{N}\sum_{j=1}^{N} \sin(\theta_j - \theta_i), \quad i = 1, \cdots, N. \tag{2.3}$$

Note that the first term in the right-hand-side of (2.3), represents the intrinsic randomness of the system, whereas the second term describes the nonlinearity of the attractive coupling. It is easy to see that the total phase $\sum_{i=1}^{N}\theta_i$ satisfies a balanced law:

$$\frac{d}{dt}\sum_{i=1}^{N}\theta_i = \sum_{i=1}^{N}\omega_i, \quad t > 0.$$

Thus, when the total sum of natural frequencies is not zero, then system (2.3) cannot have equilibria $\Theta_e = (\theta_{1e}, \cdots, \theta_{Ne})$:

$$\dot{\theta}_{ie} = 0, \quad 1 \leq i \leq N.$$

However, we may still expect existence of relative equilibria, which are the equilibria of (2.3) in a rotating coordinate frame with the angular velocity $\omega_c := \frac{1}{N}\sum_{i=1}^{N}\omega_i$. The relative equilibrium for (2.3) is called the phase-locked state. More precisely, we present its formal definition as follows.

Definition 2.1. [24, 60] *Let* $\Theta = (\theta_1, \cdots, \theta_N)$ *be a solution to* (2.3).

1. $\Theta = (\theta_1, \cdots, \theta_N)$ *is a phase-locked state if the transversal phase differences are constant along the Kuramoto flow* (2.3)*:*

$$|\theta_i(t) - \theta_j(t)| = |\theta_i(0) - \theta_j(0)|, \quad \forall\, t \geq 0, \quad 1 \leq i, j \leq N.$$

2. *The Kuramoto model* (2.3) *exhibits "complete (frequency) synchronization" asymptotically if the transversal frequencies differences approach zero asymptotically:*

$$\lim_{t\to\infty} \max_{1\leq i,j\leq N} |\dot{\theta}_i(t) - \dot{\theta}_j(t)| = 0.$$

3. *The Kuramoto model* (2.3) *exhibits "complete phase synchronization" asymptotically if the transversal phase differences approach zero asymptotically:*

$$\lim_{t\to\infty} \max_{1\leq i,j\leq N} |\theta_i(t) - \theta_j(t)| = 0.$$

4. *The Kuramoto model* (2.3) *exhibits "practical (phase) synchronization" asymptotically if the transversal phase differences satisfy*

$$\lim_{K\to\infty} \limsup_{t\to\infty} \max_{1\leq i,j\leq N} |\theta_i(t) - \theta_j(t)| = 0.$$

Remark 2.1. *1. If complete synchronization occurs asymptotically, solutions tend to phase-locked states asymptotically. We also note that for non-identical oscillators, complete phase synchronization is not possible even asymptotically. For details on the phase-locked states, we refer to [24, 56].*

2. When the average natural frequency ω_c is zero, the equilibrium solution Θ to (2.3) *which is a solution to the following system of transcendental equations:*

$$\omega_c := \sum_{i=1}^{N} \omega_i = 0, \quad \omega_i + \frac{K}{N} \sum_{j=1}^{N} \sin(\theta_j - \theta_i) = 0, \quad i = 1, \cdots, N,$$

is a phase-locked solution to (2.3) *as well.*

3. For a brief review on the classical and quantum Kuramoto type models, we refer to a recent survey paper [58].

2.1.1 Order parameters

In this part, we briefly review real-valued order parameters for the phase configuration $\Theta = (\theta_1, \cdots, \theta_N)$ and their dynamics following the presentation given in [55]. Consider the average position (centroid) of N limit-cycle oscillators $z_i = e^{\mathrm{i}\theta_i}$: for $t \geq 0$,

$$r(t)e^{\mathrm{i}\phi(t)} := \frac{1}{N} \sum_{j=1}^{N} e^{\mathrm{i}\theta_j(t)}. \tag{2.4}$$

Here we call r and ϕ the amplitude and the average phase order parameters for the N limit-cycle system, respectively. Since the right hand side of (2.4) is a convex combination of N-points on the unit circle, the amplitude $r(t)$ lies on the interval $[0, 1]$ and the cases $r = 0$ and 1 correspond to the splay state and the completely phase synchronized state, respectively. Hence, we can regard r and ϕ as quantities measuring the degree of overall synchronization and the average of phases, respectively. Note that the average phase ϕ is well-defined when $r > 0$.

We divide (2.4) by $e^{\mathrm{i}\theta_k}$ to obtain

$$re^{\mathrm{i}(\phi-\theta_k)} = \frac{1}{N}\sum_{j=1}^{N} e^{\mathrm{i}(\theta_j-\theta_k)},$$

and compare the real and imaginary parts of the above relation to find

$$r\cos(\phi-\theta_k) = \frac{1}{N}\sum_{j=1}^{N}\cos(\theta_j-\theta_k), \quad r\sin(\phi-\theta_k) = \frac{1}{N}\sum_{j=1}^{N}\sin(\theta_j-\theta_k). \tag{2.5}$$

Similarly, we divide the relation (2.4) by $e^{\mathrm{i}\phi(t)}$ and compare real and imaginary parts to see the following relations:

$$r = \frac{1}{N}\sum_{j=1}^{N}\cos(\theta_j-\phi), \quad 0 = \frac{1}{N}\sum_{j=1}^{N}\sin(\theta_j-\phi).$$

By comparing the second relation in (2.5) and the coupling terms in (2.3), it is easy to see that the Kuramoto model (2.3) can be rewritten in a mean-field form:

$$\dot{\theta}_i = \omega_i - Kr\sin(\theta_i-\phi), \quad t > 0. \tag{2.6}$$

The equation (2.6) looks decoupled, but it is coupled, because the order parameters r and ϕ are functions of other θ_j's.

We next study the dynamics of the order parameters r and ϕ. For this, we differentiate the equation (2.4) with respect to t to see

$$\dot{r}e^{\mathrm{i}\phi} + \mathrm{i}re^{\mathrm{i}\phi}\dot{\phi} = \frac{\mathrm{i}}{N}\sum_{j=1}^{N} e^{\mathrm{i}\theta_j}\dot{\theta}_j.$$

We divide the resulting equation by $e^{\mathrm{i}\phi}$ to find

$$\dot{r} + \mathrm{i}r\dot{\phi} = -\frac{1}{N}\sum_{j=1}^{N}\sin(\theta_j - \phi)\dot{\theta}_j + \frac{\mathrm{i}}{N}\sum_{j=1}^{N}\cos(\theta_j - \phi)\dot{\theta}_j. \tag{2.7}$$

We now compare the real and imaginary parts of (2.7) to obtain

$$\dot{r} = -\frac{1}{N}\sum_{j=1}^{N}\sin(\theta_j - \phi)\dot{\theta}_j, \qquad \dot{\phi} = \frac{1}{rN}\sum_{j=1}^{N}\cos(\theta_j - \phi)\dot{\theta}_j. \tag{2.8}$$

Thus, we can combine (2.6) and (2.8) to get the evolutionary system:

$$\begin{aligned} \dot{r} &= -\frac{1}{N}\sum_{j=1}^{N}\sin(\theta_j - \phi)\Big(\omega_j - Kr\sin(\theta_j - \phi)\Big), \\ \dot{\phi} &= \frac{1}{rN}\sum_{j=1}^{N}\cos(\theta_j - \phi)\Big(\omega_j - Kr\sin(\theta_j - \phi)\Big). \end{aligned} \tag{2.9}$$

Before we close this part, we present the relationship between the phase diameter $D(\Theta) := \max_{1\le i,j\le N}|\theta_i - \theta_j|$ and the order parameter r in the following proposition.

Proposition 2.1. [24] *Suppose that the phase configuration $\Theta = (\theta_1, \cdots, \theta_N)$ is confined in a half circle such that*

$$D(\Theta) < \pi.$$

Then, the following estimates hold.

1. *The order parameter r is rotationally invariant and $r \geq \cos \frac{D(\Theta)}{2}$.*

2. *The order parameter r satisfies*

$$r = 1 \quad \Longleftrightarrow \quad D(\Theta) = 0, \quad i.e., \ \ \theta_1 = \theta_2 = \cdots = \theta_N.$$

Proof. (i) For the rotational invariance of r, it suffices to show that the configuration $\Theta + \alpha \mathbf{1} := (\theta_1 + \alpha, \cdots, \theta_N + \alpha)$ and $\Theta := (\theta_1, \cdots, \theta_N)$ have the same order parameter. Let $\bar{r}$ and r be the order parameters for $\Theta + \alpha\mathbf{1}$ and Θ, respectively.

$$\bar{r} = \frac{1}{N}\Big|\sum_{j=1}^N e^{\mathrm{i}(\theta_j + \alpha)}\Big| = \frac{1}{N}\Big|e^{\mathrm{i}\alpha}\sum_{j=1}^N e^{\mathrm{i}\theta_j}\Big| = \frac{1}{N}\Big|\sum_{j=1}^N e^{\mathrm{i}\theta_j}\Big| = r.$$

(ii) Since the order parameter is rotationally invariant, without loss of generality, we may assume

$$\theta_i \in \Big(-\frac{D(\Theta)}{2}, \frac{D(\Theta)}{2}\Big), \quad 1 \leq i \leq N.$$

• Case A: We first prove that if $D(\Theta) = 0$, then $r = 1$. For this, we note that

$$\begin{aligned} r = \frac{1}{N}\Big|\sum_{j=1}^N e^{\mathrm{i}\theta_j}\Big| &= \frac{1}{N}\Big[\Big(\sum_{j=1}^N \cos\theta_j\Big)^2 + \Big(\sum_{j=1}^N \sin\theta_j\Big)^2\Big]^{\frac{1}{2}} \\ &\geq \frac{1}{N}\sum_{j=1}^N \cos\theta_j \geq \cos\frac{D(\Theta)}{2}. \end{aligned}$$

Thus, if $D(\Theta) = 0$, then

$$1 \geq r \geq \cos\frac{D(\Theta)}{2} = 1, \quad \text{i.e.,} \quad r = 1.$$

• Case B: Note that

$$z_i = e^{\mathrm{i}\theta_i} \in \mathbb{S}^1 \subset \mathbb{C}, \qquad |z_1 + \cdots + z_N| = Nr,$$

Suppose that $r = 1$, then we have

$$|z_1 + \cdots + z_N| = N. \tag{2.10}$$

We now claim:

$$\theta_i = \theta_j, \quad 1 \le i, j \le N.$$

It follows from the relation (2.10) that we have

$$0 = |z_1 + \cdots + z_N|^2 - N^2 = \sum_{i=1}^{N} |z_i|^2 - N^2 + 2 \sum_{1 \le i < j \le N} \cos(\theta_i - \theta_j).$$

This yields

$$\sum_{1 \le i < j \le N} \cos(\theta_i - \theta_j) = \frac{N(N-1)}{2}.$$

This again implies

$$\cos(\theta_i - \theta_j) = 1 \quad \text{i.e.,} \quad \theta_i - \theta_j = 0, \quad 1 \le i, j \le N.$$

□

2.1.2 A gradient flow formulation

Note that the right hand side of (2.3) is 2π-periodic, so the system (2.3) is a dynamical system on N-tori $\mathbb{T}^N$. However, for the description of a gradient flow formulation, we lift the system (2.3) to a dynamical system on $\mathbb{R}^N$ by a straightforward lifting. So the trajectory of $\Theta = (\theta_1, \cdots, \theta_N)$ is not necessarily bounded as a subset of $\mathbb{R}^N$. In [92], from an analogy with the XY-model in statistical physics, Hemmen and Wreszinski observed that the Kuramoto model (2.3) can be formulated as a gradient flow with an analytic potential. More precisely, they introduced the analytic potential V_p:

$$V_p(\Theta) := -\sum_{i=1}^{N} \omega_i \theta_i + \frac{K}{2N} \sum_{i,j=1}^{N} \Big(1 - \cos(\theta_j - \theta_i) \Big). \tag{2.11}$$

By direct calculation, it is easy to see that the Kuramoto model (2.3) can be rewritten as a gradient flow form:

$$\dot{\Theta}(t) = -\nabla V_p(\Theta). \tag{2.12}$$

Note that the potential function V_p is neither convex nor bounded below a priori. This gradient formulation has been crucially used to prove the emergence of phase-locked states from generic initial configurations [35, 57].

2.1.3 Emergent dynamics

In this part, we recall the uniform boundedness of fluctuations of phases around the averaged motion, as well as mass concentration of the identical oscillators around the average phase. Since we regard the system (2.12) as a dynamical system on $\mathbb{R}^N$, the uniform boundedness of fluctuations around the average phase motion is not clear a priori. In [57], authors showed that the relative phases are uniformly bounded, if more than half of oscillators are confined in a small arc and the coupling strength is sufficiently large.

Proposition 2.2. [57] *Suppose that the initial configuration Θ_0 satisfies*

$$\theta_{j0} \in [-\pi, \pi), \quad 1 \le j \le N,$$

and let n_0, l, and K satisfy

$$n_0 \in \mathbb{Z}_+ \cap \Big(\frac{N}{2}, N\Big], \quad l \in \Big(0, 2\cos^{-1}\frac{N-n_0}{n_0}\Big),$$

$$\max_{1 \le j,k \le n_0} |\theta_{j0} - \theta_{k0}| < l, \quad K > \frac{\max_{i,j} |\omega_i - \omega_j|}{\frac{n_0}{N}\sin l - \frac{2(N-n_0)}{N}\sin\frac{l}{2}}.$$

Let Θ be a global solution to the system (2.3)*. Then, we have*

$$\sup_{0 \le t < \infty} \max_{i,j} |\theta_i(t) - \theta_j(t)| \le 4\pi + l.$$

Remark 2.2. *For identical oscillators with $\omega_i = \omega$, Dong and Xue [35] used the gradient flow formulation* (2.12) *to show that for all initial configurations and positive coupling strength, the Kuramoto flow* (2.3) *tends to phase-locked states. Moreover, the uniform boundedness of Proposition 2.1 and gradient flow formulation yields the formation of phase-locked states.*

Next, we consider the Kuramoto model for identical oscillators:

$$\dot{\theta}_i = \omega + \frac{K}{N}\sum_{j=1}^{N} \sin(\theta_j - \theta_i), \quad i = 1, \cdots, N. \tag{2.13}$$

For a dynamical solution $\Theta(t)$ to (2.13), we divide the oscillator set $\mathcal{N} := \{1, \cdots, N\}$ into synchronous and anti-synchronous oscillators, with respect to the overall phase $\phi(t) := \omega t$:

$$\mathcal{I}_s := \{j : \lim_{t\to\infty} |\theta_j(t) - \phi(t)| = 0\}, \qquad \mathcal{I}_b := \{j : \lim_{t\to\infty} |\theta_j(t) - \phi(t)| = \pi\}.$$

Proposition 2.3. [57] *Let $\Theta = (\theta_1, \cdots, \theta_N)$ be a solution to* (2.13) *with initial configuration Θ_0 satisfying the following conditions:*

$$\sum_{i=1}^{N} \Omega_i = 0, \quad r_0 > 0, \qquad \theta_{k0} \ne \theta_{j0}, \qquad 1 \le k, j \le N.$$

Then, $\{\mathcal{I}_s, \mathcal{I}_b\}$ *is a partition of* $\mathcal{N}$ *and we have*

$$|\mathcal{I}_b| \le 1,$$

where $|A|$ *denotes the cardinality of the set* A*.*

Remark 2.3. *1. The state with* $|\mathcal{I}_b| = 1$ *is called the bi-polar state, which is well known to be unstable [24].*

2. For a detailed survey on the Kuramoto model (2.3)*, we refer to [1, 9, 24, 25, 35, 36, 37, 57, 58, 59, 65]).*

2.2 The Kuramoto-Sakaguchi (K-S) equation

In this subsection, we discuss the kinetic counterpart of (2.3). Consider a situation where the number of oscillators, which we denote by N in (2.3) goes to infinity. In this mean-field limit, it is more convenient to rewrite the system (2.3), as a dynamical system on the phase space $\mathbb{T} \times \mathbb{R}$, $\mathbb{T} := \mathbb{R}/2\pi\mathbb{Z} = [0, 2\pi]$ for (θ, ω):

$$\begin{cases} \dfrac{d\theta_i}{dt} = \omega_i - \dfrac{K}{N}\displaystyle\sum_{j=1}^{N} \sin(\theta_i - \theta_j), \\ \dfrac{d\omega_i}{dt} = 0, \quad t > 0. \end{cases} \tag{2.14}$$

Since we are dealing with a large oscillator system $N \gg 1$, we can introduce a probability density function $f = f(\theta, \omega, t)$ to approximate the N-oscillator system (2.14). Based on the standard BBGKY Hierarchy argument [68], we can derive the K-S equation:

$$\begin{aligned} &\partial_t f + \partial_\theta(v[f]f) = 0, \qquad (\theta,\omega) \in \mathbb{T}\times\mathbb{R}, \quad t>0, \\ &v[f](\theta,\omega,t) = \omega - KF[\rho], \quad F[\rho] := \int_{\mathbb{T}} \sin(\theta-\theta_*)\rho(\theta_*,t)\ d\theta_*. \end{aligned} \tag{2.15}$$

where $\rho = \rho(\theta,t)$ is the local mass density function, which corresponds to the θ-marginal density function of f:

$$\rho(\theta,t) := \int_{\mathbb{R}} f(\theta,\omega,t)\ d\omega, \quad t \geq 0.$$

Remark 2.4. *The rigorous derivation from* (2.3) *to* (2.15) *was done by Lancellotti [71] using Neunzert's method [81]. In [71, Theorem 2 and Remark 3], Lancellotti showed that there exists a unique classical solution to* (2.15)*, whenever the initial datum is C^1.*

We next recall some conservation laws for the K-S equation.

Lemma 2.1. [2] *Let $f = f(\theta,\omega,t)$ be a $\mathcal{C}^1$-solution to* (2.15)*, with initial datum f_0 satisfying the following conditions:*

$$\int_{\mathbb{T}} f_0\ d\theta = g(\omega), \quad \iint_{\mathbb{T}\times\mathbb{R}} f_0\ d\theta d\omega = 1.$$

Then, we have

$$\int_{\mathbb{T}} f(\theta,\omega)\ d\theta = g(\omega), \quad \int_{\mathbb{T}\times\mathbb{R}} f\ d\theta d\omega = 1, \quad t \geq 0.$$

Proof. The identities follow from a direct computation as follows.

$$\frac{d}{dt}\int_{\mathbb{T}} f\ d\theta = \int_{\mathbb{T}} \partial_t f\ d\theta = -\int_{\mathbb{T}} \partial_\theta(v[f]f)\ d\theta = 0,$$

and

$$\frac{d}{dt}\iint_{\mathbb{T}\times\mathbb{R}} f\ d\theta d\omega = \int_{\mathbb{R}} \frac{d}{dt}\Big(\int_{\mathbb{T}} f\ d\theta\Big) d\omega = 0.$$

$\square$

Remark 2.5. *Lemma 2.1 yields that for any test function $h = h(\omega)$, we have*

$$\iint_{\mathbb{T}\times\mathbb{R}} h(\omega) f\ d\theta d\omega = \int_{\mathbb{R}} h(\omega)\Big(\int_{\mathbb{T}} f\ d\theta\Big) d\omega = \int_{\mathbb{R}} hg\ d\omega = \int_{\mathbb{T}\times\mathbb{R}} h(\omega) f_0\ d\theta d\omega.$$

2.2.1 Order parameters

In this part, we introduce real-valued order parameters $R = R(t)$ and $\phi = \phi(t)$ which measure the overall degree of synchronization for the K-S equation (2.15). Such order parameters will be used to simplify the expression of the system (2.15). As a straightforward generalization of the order parameters (2.4) for the Kuramoto model, we define real order parameters R and ϕ for the K-S equation [1, 55]:

$$R(t)e^{\mathrm{i}\phi(t)} := \iint_{\mathbb{T}\times\mathbb{R}} e^{\mathrm{i}\theta} f\ d\theta d\omega = \int_{\mathbb{T}} e^{\mathrm{i}\theta} \rho\ d\theta, \quad t \geq 0. \tag{2.16}$$

To avoid the confusion with the amplitude order parameter r for (2.4), we use a capital letter R instead of r for the K-S equation. As for the Kuramoto model, the average phase ϕ is well-defined when $R > 0$. We divide (2.16) by $e^{i\phi}$ on both sides to get

$$R(t) = \int_{\mathbb{T}} e^{\mathrm{i}(\theta-\phi(t))} \rho\ d\theta = \int_{\mathbb{T}} \langle e^{\mathrm{i}\theta}, e^{\mathrm{i}\phi(t)}\rangle \rho\ d\theta. \tag{2.17}$$

By comparing real and imaginary part on both sides of (2.17), we obtain

$$R(t) = \int_{\mathbb{T}} \cos(\theta - \phi(t))\rho(\theta, t)\ d\theta \qquad \text{and} \qquad 0 = \int_{\mathbb{T}} \sin(\theta - \phi(t))\rho(\theta, t)\ d\theta. \tag{2.18}$$

On the other hand, we use (2.18) to rewrite the linear operator $F[\rho]$ in terms of order parameters:

$$\begin{aligned}
F[\rho] &= \int_{\mathbb{T}} \sin(\theta - \theta_*)\rho(\theta_*, t)\ d\theta_* \\
&= \int_{\mathbb{T}} \sin\big((\theta - \phi) - (\theta_* - \phi)\big)\rho(\theta_*, t)\ d\theta_* \\
&= \int_{\mathbb{T}} \big(\sin(\theta - \phi)\cos(\theta_* - \phi) - \cos(\theta - \phi)\sin(\theta_* - \phi)\big)\rho(\theta_*, t)\ d\theta_* \\
&= \sin(\theta - \phi)\int_{\mathbb{T}} \cos(\theta_* - \phi)\rho(\theta_*, t)\ d\theta_* \\
&\quad - \cos(\theta - \phi)\int_{\mathbb{T}} \sin(\theta_* - \phi)\rho(\theta_*, t)\ d\theta_* \\
&= R\sin(\theta - \phi).
\end{aligned} \tag{2.19}$$

Thus, we can rewrite the Kuramoto-Sakaguchi equation (2.15) as an equivalent form:

$$\partial_t f + \partial_\theta\Big(\omega f - KR\sin(\theta - \phi)f\Big) = 0. \tag{2.20}$$

For the identical oscillator case with $g(\Omega) = \delta(\omega)$, by integrating (2.20) with respect to ω we obtain

$$\partial_t \rho + \partial_\theta\Big(\omega \rho - KR\sin(\theta - \phi)\rho\Big) = 0. \tag{2.21}$$

2.2.2 A gradient flow formulation

In this part, we discuss the gradient flow formulation of the K-S equation with $g(\omega) = \delta(\omega)$, i.e., we will write the system (2.3) as a gradient flow in the probability space $\mathbb{P}(\mathbb{T})$ equipped with the Wasserstein metric W_2. For this, we adapt the differential calculus introduced by Felix Otto [83], which has proven to be a powerful tool for the study of the dynamics and stability of evolutionary equations (we refer the reader to [66, 82, 83] for the pioneering works on this topic) (see Appendix A). We first set up a potential function for our gradient flow. In analogy with (2.11) for the Kuramoto model, it is natural to come up with the following potential function for the K-S equation (4.1).

$$V_k(\rho(t)) := \frac{K}{2} \iint_{\mathbb{T}^2} (1 - \cos(\theta_* - \theta))\rho(\theta_*, t)\rho(\theta, t)\ d\theta_* d\theta.$$

Here the subscript k stands for "*kinetic*". We next present another handy expression for $V_k(\rho)$ in terms of the order parameter R given in (2.16). First, we let $\sigma : \mathbb{T} \to \mathbb{R}^2$ be defined by $\sigma(\theta) := (\cos\theta, \sin\theta)$. Then, given ρ in $\mathbb{P}(\mathbb{T})$, the above potential function V_k can be written as

$$V_k(\rho(t)) = \frac{K}{2} - \frac{K}{2} \iint_{\mathbb{T}^2} \sigma(\theta) \cdot \sigma(\theta_*)\rho(\theta, t)\rho(\theta_*, t)\ d\theta_* d\theta, \tag{2.22}$$

where $\sigma(\theta) \cdot \sigma(\theta_*)$ denotes the inner product between the two vectors in $\mathbb{R}^2$. We define the total momentum of ρ by

$$J := \int_{\mathbb{T}} \sigma(\theta)\rho(\theta)\ d\theta \in \mathbb{R}^2. \tag{2.23}$$

When we regard $Re^{i\phi}$ in (2.16) as a vector in $\mathbb{R}^2$, we have

$$J = Re^{\mathrm{i}\phi}.$$

From this, we can write

$$V_k(\rho) = \frac{K}{2}(1 - |J|^2), \quad \text{equivalently} \quad V_k(\rho) = \frac{K}{2}(1 - R^2). \tag{2.24}$$

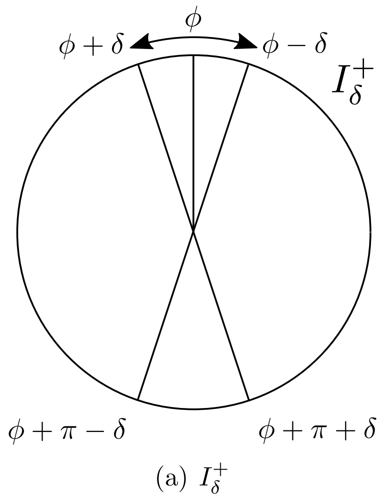


Figure 5.1: Geometric descriptions of I_δ^+ and I_δ^-

By straightforward application of Otto calculus (see Appendix A), we have

$$\operatorname{grad}_\rho V_k = -K\nabla(\sigma(\theta)\cdot J) = KR\sin(\theta-\phi),$$

where grad_ρ denotes the gradient with respect to the Wasserstein metric. We now substitute this in (A.2) to get the gradient flow of V_k as the one-parameter family $t\in[0,\varepsilon)\mapsto\rho_t\in\mathbb{P}(\mathbb{T})$ satisfying

$$\partial_t\rho = \partial_\theta\left(\rho\operatorname{grad}_\rho V_k\right) = \partial_\theta\left(\rho KR\sin(\theta-\phi)\right).$$

This is the same as (2.21) when $\omega=0$, which verifies the gradient flow structure of the K-S equation.

3 Discussion of the main results

In this section, we briefly present our three main results on the emergent dynamics of the K-S equation. The proofs will be given in the following three sections.

3.1 Emergence of point attractors

In this subsequent, we briefly discuss how point attractors can emerge from generic smooth initial data, asymptotically along the Kuramoto-Sakaguchi flow with $g = \delta$. For the finite-dimensional Kuramoto model for N-identical oscillators, if we start from generic initial datum

$$r_0 = \frac{1}{N}\Big|\sum_{j=1}^{N} e^{\mathrm{i}\theta_{j0}}\Big| > 0, \qquad \theta_{i0} \neq \theta_{j0}, \quad 1 \leq i, j \leq N,$$

then, any positive coupling strength K will push the initial configuration to two possible scenarios. First scenario to happen is that the phases will concentrate near the time-dependent average phase which is also dynamic along the Kuramoto flow, and the second scenario will be a bi-polar configuration where $(N-1)$ oscillators will aggregate toward the average phase, and the remaining one oscillator will approach the antipodal phase of the average phase (see Proposition 2.2). By an easy perturbation argument, we see that the bi-polar configuration is unstable (see [24]). Thus, the completely phase synchronized state which we call one-point attractor in this paper, is the only stable asymptotic state for the Kuramoto model in any positive coupling regime. Thus, it is interesting to figure out what will happen for the K-S equation which can be obtained from the Kuramoto model as $N \to \infty$:

> Does the same asymptotic patterns as in the Kuramoto flow emerge for the K-S flow from generic initial configurations?

In fact, we will argue that the unstable bi-polar configurations can not be reachable from a generic smooth initial datum along the K-S flow. To justify and quantify our claim, we will first define a set consisting of two disjoint intervals containing the average phase and its antipodal point, respectively, and then we will show that the mass outside this set will approach to zero asymptotically. Finally, we will argue that the mass around the antipodal point of the average phase will tend to zero exponentially, i.e., all mass will aggregate in any small neighborhood of the average phase. Hence, asymptotically, one-point attractor will emerge. In processing the above procedures, analysis on the dynamics of R and ϕ will play a key role. To quantify the above sketched arguments, we define for a positive constant $\delta > 0$, a set $I_\delta := I_\delta^+ \cup I_\delta^-$ (see Figure 5.1):

$$\begin{aligned} I_\delta^+(t) &:= \{\theta \in \mathbb{T} : |\theta - \phi(t)| < \delta\}, \\ I_\delta^-(t) &:= \{\theta \in \mathbb{T} : |\theta - (\phi(t) + \pi)| < \delta\}. \end{aligned} \tag{3.1}$$

Consequently, we will see that eventually all the mass of ρ will be concentrated in the region $I_\delta^+ \cup I_\delta^-$ for every $\delta > 0$. Since R is monotonic, if $R_0 > 0$, we necessarily have

$$\int_{I_\delta^+} \rho \, d\theta > \int_{I_\delta^-} \rho \, d\theta.$$

Then, it is reasonable to expect that the mass will exit I_δ^- and enter the interval I_δ^+, and eventually will yield a point attractor, provided that we can control the rotation of $\phi(t)$. This is proved in (4.14) and (4.15), whereas in Proposition 4.1 we can control the rotation of $\phi(t)$, using the fact that $\dot{R}(t) \to 0$.

We are now ready to state our first result on the emergence of point attractors.

Theorem 3.1. *Suppose that the coupling strength, the density function g and the initial datum satisfy*

$$K > 0, \quad g(\omega) = \delta(\omega), \quad \rho_0 \in C^1, \quad \int_{\mathbb{T}} \rho_0 \, d\theta = 1 \quad \text{and} \quad R_0 := R(0) > 0.$$

Then, for a classical solution $\rho : \mathbb{T} \times \mathbb{R}_+ \to \mathbb{R}$ to (2.21)*, the mass concentrates around the average phase $\phi(t)$ asymptotically. More precisely, for any $\delta > 0$, there exists $T_1 = T_1(\delta) \geq 0$, such that*

$$\lim_{t \to \infty} \int_{I_\delta^+(t)} \rho(\theta, t) \, d\theta = 1 \quad \text{and}$$
$$\int_{I_\delta^-(t)} |\rho(\theta, t)|^2 \, d\theta \leq e^{-\frac{R(0)(\cos\delta)}{2} K(t - T_1)} \int_{I_\delta^-(T_1)} |\rho(\theta, T_1)|^2 \, d\theta \qquad \forall \, t \geq T_1.$$

Proof. The proof will be given in Section 4.2. □

Remark 3.1. *Concentration of mass around ϕ has been proved in [14] by a different argument without the exponential decay estimate of mass in the interval I_δ^-.*

3.2 Emergence of phase concentration

In this subsection and the next, we will show that there exists an interval centered at ϕ where the mass will concentrate asymptotically, when the coupling strength is sufficiently large. Moreover, we will also present a lower bound for the asymptotic amplitude order parameter $R = R(t)$ which tends to unity as $K \to \infty$. Before we discuss our second and third main results, we first recall corresponding results for the Kuramoto model in a large coupling strength regime (see Proposition 2.2). Consider an ensemble consisting of N Kuramoto oscillators and assume that more than $\frac{N}{2}$ oscillators are confined in a small interval $\mathcal{I}$ at some instant. We divide the ensemble into two sub-ensembles consisting of confining oscillators in the interval $\mathcal{I}$ and the rest of it. In this situation, if we choose a sufficiently large coupling strength, then the confining oscillators will stay in the interval and drifting oscillators might enter neighboring copies of the interval $(\mathcal{I}+2\pi)\cup(\mathcal{I}-2\pi)$. Thus, trajectories of oscillators will be bounded eventually. We now return to the K-S equation. As we have seen in the previous subsection, for identical oscillators, the total mass will concentrate asymptotically on the average phase. Thus, for the density function $g = g(\omega)$ with a compact support, we can imagine that a similar scenario for the Kuramoto model will happen, i.e., once we choose a large coupling strength compared to the size of the support of the natural frequency density function g, we can guess that mass will be confined inside some small interval around the average phase. In fact, this is the case. To be more precise, assume that $g = g(\omega)$ is compactly supported inside the interval $[-M, M]$ and some significant portion of local mass of f_0 is concentrated on some time-dependent interval $L(t)$ centered the average phase $\phi(t)$: for some positive constant $C > 0$

$$\iint_{L(0)\times\mathbb{R}} f_0(\theta,\omega)\ d\theta d\omega > C.$$

In this setting, for a large coupling strength $K \gg M$, we will derive

$$\frac{d}{dt}\iint_{L(t)\times\mathbb{R}} f(\theta,\omega,t)\ d\theta d\omega \ge 0,$$
$$\frac{d}{dt}\int_{L(t)} |f(\theta,\omega,t)|^2\ d\theta \ge K|\mathcal{O}(1)|\int_{L(t)} |f(\theta,\omega,t)|^2\ d\theta, \quad \omega\in \text{supp}\ g(\omega).$$

Thus, we have

$$\frac{d}{dt}\iint_{L(t)\times\mathbb{R}} f(\theta,\omega,t)\ d\theta d\omega \ge 0, \qquad \int_{L(t)} |f(\theta,\omega,t)|^2\ d\theta \ge Ce^{Ct},$$
$$\omega\in \text{supp}\ g(\omega).$$

Note that the first estimate says that the mass on the set $L(t)\times\mathbb{R}$ in the phase space is nondecreasing, i.e., the mass does not leak to complement of this set, and the second estimate tells us that for each fixed $\omega \in \text{supp}\ g(\omega)$, there should be some mass concentration. In this sense, we may say that the time-dependent set $L(t)\times\mathbb{R}$ converges to an invariant manifold for the K-S flow. For a precise statement, we set

$$\mathcal{M}_*(\varepsilon_0,\gamma_0) := \frac{2+\varepsilon_0+\cos\gamma_0}{(1+\sin\gamma_0)(1+\cos\gamma_0)}. \tag{3.2}$$

We now ready to state our second main result summarizing the above arguments.

Theorem 3.2. *Suppose that the following conditions hold.*

1. *The frequency density function $g=g(\omega)$ and coupling strength K satisfies*

$$\text{supp}\, g(\omega)\subset(-M,M), \quad K > \frac{M}{\varepsilon_0}\left(1+\frac{1}{\varepsilon_0}\right), \quad \varepsilon_0 \in \left(0, \frac{3\sqrt{3}}{4}-1\right).$$

2. Suppose that initial datum f_0 satisfies

$$(i)\ f_0(\theta,\omega) = 0 \quad \text{in} \quad \mathbb{T} \times (\mathbb{R}\backslash[-M,M]), \quad ||f_0||_{L^\infty} < \infty,$$
$$(ii)\ \iint_{L^+_{\gamma_0}(0)\times\mathbb{R}} f(\theta,\omega,0)\ d\theta d\omega \geq \mathcal{M}_*(\varepsilon_0,\gamma_0),$$

where γ_0 satisfies

$$\frac{\pi}{3} \leq \gamma_0 < \arcsin\Big(1 - \frac{2\varepsilon_0}{2\sqrt{3}+1}\Big).$$

Then, for any C^1-solution to (2.15)*, there exists a time-dependent interval $L(t) \subset \mathbb{T}$ centered around $\phi(t)$, with fixed width such that*

$$\frac{d}{dt}\iint_{L(t)\times\mathbb{R}} f(\theta,\omega,t)\ d\theta d\omega \geq 0, \qquad \int_{L(t)} |f(\theta,\omega,t)|^2\ d\theta \geq Ce^{Ct},$$
$$\omega \in supp\ g(\omega).$$

Proof. We present its proof in Section 5. □

3.3 Asymptotic dynamics of the order parameter

Notice that in Theorem 3.2, we assumed a certain lower bound on the mass in a certain interval. We remove such assumption for our third main result which we now describe. For the Kuramoto model (2.3), the dynamics of the order parameter r does play a key role in the recent resolution of the complete synchronization problem in [57]. As noticed in Proposition 2.1, for the Kuramoto model, if the order parameter r is close to 1, we can say that the configuration is close to complete phase synchronization where all phases are concentrated at some common phase. Our third result is concerned about the estimation of the order parameter in a large coupling strength regime. More precisely, we will obtain a positive lower bound as

$$\liminf_{t\to\infty} R(t) \geq 1 - \frac{|\mathcal{O}(1)|}{\sqrt{K}}, \quad \text{for } K \gg 1.$$

This certainly implies that as $K \to \infty$,

$$\lim_{K\to\infty} \liminf_{t\to\infty} R(t) = 1.$$

By Proposition 2.1, this means the formation of complete phase synchronization in $K \to \infty$ limit, which can be understood as an emergence of practical synchronization. Below, we state our third result.

Theorem 3.3. *Let* $f : \mathbb{T} \times \mathbb{R} \times \mathbb{R}^+ \to \mathbb{R}$ *be a classical solution to* (2.15)*. Suppose* g *is supported on the interval* $[-M, M]$*,* $R(0) := R_0 > 0$*, and* K *is sufficiently large (depending on the support of* g *and* $1/R_0$*). Then,*

$$\liminf_{t\to\infty} R(t) \geq R_\infty := 1 + \frac{M}{K} - \sqrt{\frac{M^2}{K^2} + 4\frac{M}{K}}$$

and

$$\lim_{t\to\infty} ||f_t \mathbb{1}_{\mathbb{T}\setminus L_\infty(t)}||_{L^\infty(\mathbb{T}\times\mathbb{R})} = 0.$$

Here, $L_\infty(t) \subset \mathbb{T}$ *is a time dependent interval, centered at* $\phi(t)$ *with the constant width*

$$\arccos\left(\sqrt{1 - \left[\frac{M}{K}\frac{(1+R_\infty)}{R_\infty^2} + \frac{1-R_\infty}{R_\infty}\right]^2}\right) + \varepsilon,$$

where ε *is an arbitrary constant in* $(0,1)$*. Notice that as* $K \to \infty$ *the width of* L_∞ *can be made arbitrarily small and* R_∞ *tends to* 1.

Proof. The proof will be given in Section 6.3. □

In the following three sections, we will present proofs of the main theorems and of many lemmata.

4 Emergence of point attractors

In this section, we present existence of point attractors for the K-S equation with $g = \delta$ from a generic initial datum using the time-asymptotic approach. Without loss of generality, we may assume that the common natural frequency ω_c is zero and that the corresponding density function $g = g(\omega)$ satisfies

$$g(\omega) = \delta_0.$$

We first note that the local mass density ρ satisfies the following two equivalent equations:

$$\begin{cases} \partial_t \rho + \partial_\theta (v[\rho]\rho) = 0, \quad \theta \in \mathbb{T}, \quad t > 0, \\ v[\rho](\theta, t) = -K \displaystyle\int_{\mathbb{T}} \sin(\theta - \theta_*) \rho(\theta_*, t) \, d\theta_*, \end{cases} \tag{4.1}$$

or equivalently

$$\partial_t \rho - \partial_\theta \Big(KR\rho \sin(\theta - \phi) \Big) = 0.$$

In the following two subsections, we will study the dynamics of the amplitude order parameter R and present the proof of Theorem 3.1.

4.1 Dynamics of order parameters

In this subsection, we derive a coupled dynamical system for the order parameters R and ϕ, and using this dynamical system, we analyze their asymptotics. Note that the complete phase synchronization occurs if and only if $R \to 1$ as $t \to \infty$.

For the derivation of dynamics of R and ϕ, we differentiate the defining relation (2.16) with respect to t, and we obtain

$$\dot{R}e^{\mathrm{i}\phi} + \mathrm{i}R\dot{\phi}e^{\mathrm{i}\phi} = \int_{\mathbb{T}} \partial_t \rho(\theta, t) e^{\mathrm{i}\theta}\ d\theta. \tag{4.2}$$

We divide (4.2) by $e^{i\phi}$ on both sides to get

$$\dot{R} + \mathrm{i}R\dot{\phi} = \int_{\mathbb{T}} \partial_t \rho(\theta, t) e^{\mathrm{i}(\theta-\phi)}\ d\theta. \tag{4.3}$$

We compare real and imaginary parts of (4.3) and employ (4.1) to derive relations for R and ϕ:

$$\begin{aligned} \dot{R} &= \int_{\mathbb{T}} \partial_t \rho(\theta, t) \cos(\theta - \phi)\ d\theta, \quad t > 0, \\ \dot{\phi} &= \frac{1}{R} \int_{\mathbb{T}} \partial_t \rho(\theta, t) \sin(\theta - \phi)\ d\theta. \end{aligned} \tag{4.4}$$

Lemma 4.1. *Let ρ be a solution to* (4.1) *and let R and ϕ be the order parameters defined by the relation* (2.17)*. Then, R and ϕ satisfy*

(i) $\dot{R} = KR \int_{\mathbb{T}} \sin^2(\theta - \phi)\rho(\theta, t)\ d\theta, \quad \dot{\phi} = -\frac{K}{2} \int_{\mathbb{T}} \sin\left(2(\theta - \phi)\right) \rho(\theta, t)\ d\theta.$

(ii) $\ddot{R} = \frac{(\dot{R})^2}{R} + 2R(\dot{\phi})^2 - 2(KR)^2 \int_{\mathbb{T}} \sin^2(\theta - \phi) \cos(\theta - \phi) \rho(\theta, t)\ d\theta.$

Proof. (i) We use (4.4) and (4.1) to obtain

$$\begin{aligned}\dot{R} &= \int_{\mathbb{T}} \cos(\theta - \phi)\partial_t \rho \, d\theta \\ &= KR \int_{\mathbb{T}} \cos(\theta - \phi)\partial_\theta \Big[\rho(\theta, t)\sin(\theta - \phi)\Big] \, d\theta \\ &= KR \int_{\mathbb{T}} \sin^2(\theta - \phi)\rho(\theta, t) \, d\theta.\end{aligned} \tag{4.5}$$

Similarly, we have

$$\begin{aligned}\dot{\phi} &= \frac{1}{R}\int_{\mathbb{T}} \sin(\theta - \phi)\partial_t \rho(\theta, t) \, d\theta \\ &= K \int_{\mathbb{T}} \sin(\theta - \phi)\partial_\theta \Big[\rho(\theta, t)\sin(\theta - \phi)\Big] \, d\theta \\ &= -K \int_{\mathbb{T}} \sin(\theta - \phi)\cos(\theta - \phi)\rho(\theta, t) \, d\theta \\ &= -\frac{K}{2}\int_{\mathbb{T}} \sin\left(2(\theta - \phi)\right)\rho(\theta, t) \, d\theta.\end{aligned}$$

(ii) We again differentiate (4.5) with respect to t to get

$$\begin{aligned}\ddot{R} &= K\dot{R}\int_{\mathbb{T}} \sin^2(\theta - \phi)\rho(\theta, t) \, d\theta - 2KR\dot{\phi}\int_{\mathbb{T}} \sin(\theta - \phi)\cos(\theta - \phi)\rho(\theta, t) \, d\theta \\ &\quad + KR\int_{\mathbb{T}} \sin^2(\theta - \phi)\partial_t \rho(\theta, t) \, d\theta \\ &= \frac{(\dot{R})^2}{R} + 2R(\dot{\phi})^2 + (KR)^2\int_{\mathbb{T}} \sin^2(\theta - \phi)\partial_\theta\Big[\rho(\theta, t)\sin(\theta - \phi)\Big] \, d\theta \\ &= \frac{(\dot{R})^2}{R} + 2R(\dot{\phi})^2 - (KR)^2\int_{\mathbb{T}} 2\sin^2(\theta - \phi)\cos(\theta - \phi)\rho(\theta, t) \, d\theta.\end{aligned}$$

This yields the desired result. □

Based on the dynamics given in Lemma 4.1, we study asymptotics of R and ϕ.

Proposition 4.1. *Let $\rho = \rho(\theta, t)$ be a solution to* (4.1) *with the initial datum ρ_0 satisfying*

$$R_0 > 0 \quad \textit{and} \quad \rho_0 \in \mathbb{P}(\mathbb{T}) \cap C^1(\mathbb{T}).$$

Then, there exists a positive constant $R_\infty \leq 1$ such that

$$\begin{aligned}
&(i) \quad \inf_{0 \leq t < \infty} R(t) \geq R_0 > 0, \quad \lim_{t \to \infty} (R(t), \dot{R}(t)) = (R_\infty, 0), \\
&(ii) \ |\dot{\phi}(t)| \leq K(1 - R(t)), \quad t \geq 0, \quad \lim_{t \to \infty} |\dot{\phi}(t)| = 0.
\end{aligned}$$

Proof. (i) Note that estimates in Lemma 4.1 yield the uniform boundedness of $\dot{R}, \ddot{R}$ and $\dot{\phi}$. So $R, \dot{R}$ and ϕ are Lipschitz continuous. Moreover, we have

$$\dot{R} \geq 0, \quad \text{thus,} \quad R(t) \geq R_0 \quad \forall t \in [0, \infty). \tag{4.6}$$

On the other hand, since $R \leq 1$, R must converge to $R_\infty \leq 1$.

Suppose $\dot{R}$ does not converge to zero. Since $\dot{R} \geq 0$, we can find a sequence of time $\{t_n\}$ such that $t_n \uparrow \infty$ as $n \to \infty$ and $\dot{R}(t_n) > \alpha$ for some positive constant α. From Lemma 4.1, we attain the Lipschitz continuity of $\dot{R}$ with $|\ddot{R}| \leq \frac{K^2}{R_0} + K + 2K^2 =: C_0$, which yieds

$$\dot{R}(s) \geq \dot{R}(t_n) - C_0|t_n - s| \geq \alpha - \frac{\alpha}{2} = \frac{\alpha}{2}, \quad \forall s \in (t_n - \frac{\alpha}{2C_0}, t_n + \frac{\alpha}{2C_0}).$$

Thus, we have

$$\int_{t_n - \frac{\alpha}{2C_0}}^{\infty} \dot{R}(s)\, ds \geq \int_{t_n - \frac{\alpha}{2C_0}}^{t_n + \frac{\alpha}{2C_0}} \dot{R}(s)\, ds \geq \frac{\alpha^2}{2C_0}.$$

This contradicts the convergence of R, i.e.,

$$\lim_{a \to \infty} \int_a^{\infty} \dot{R}(s)\, ds = 0.$$

Hence, we attain $\dot{R} \to 0$ as $t \to \infty$.

(ii) We next derive the estimate:

$$- K(1 - R) \leq \dot{\phi} \leq K(1 - R). \tag{4.7}$$

For the second inequality, we use the second result in Lemma 4.1 to obtain

$$\begin{aligned}
\dot{\phi} &= -K \int_{\mathbb{T}} \sin(\theta - \phi) \cos(\theta - \phi) \rho(\theta, t) \, d\theta \\
&= -K \int_{\mathbb{T}} \underbrace{\big(\sin(\theta - \phi) - 1 \big) \big(\cos(\theta - \phi) - 1 \big)}_{\geq 0} \rho(\theta, t) \, d\theta \\
&- K \int_{\mathbb{T}} \big(\cos(\theta - \phi) + \sin(\theta - \phi) - 1 \big) \rho(\theta, t) \, d\theta \\
&\leq -K \int_{\mathbb{T}} \cos(\theta - \phi) \rho(\theta, t) \, d\theta - K \int_{\mathbb{T}} \sin(\theta - \phi) \rho(\theta, t) \, d\theta + K \\
&= -KR + K = K(1 - R).
\end{aligned}$$

In the last line we used (2.18). Similarly, we get the first inequality in (4.7). For the remaining estimate, we use the formulas for $\dot{\phi}$ and $\dot{R}$ in Lemma 4.1, the monotonicity of R, and the Cauchy-Schwarz inequality to get

$$\begin{aligned}
|\, \dot{\phi} \,| &\leq K \int_{\mathbb{T}} |\sin(\theta - \phi)| \, \rho \, d\theta \leq K \left(\int_{\mathbb{T}} |\sin(\theta - \phi)|^2 \, \rho \, d\theta \right)^{\frac{1}{2}} \\
&\leq \sqrt{\frac{K}{R}} \sqrt{|\dot{R}|} \leq \sqrt{\frac{K}{R(0)}} \sqrt{|\dot{R}|}, \text{ (since } R \geq R_0 \text{ from (i)).}
\end{aligned} \tag{4.8}$$

Since $\dot{R} \to 0$ (see item (i)), we conclude

$$\lim_{t \to \infty} |\dot{\phi}(t)| = 0.$$

□

We are now ready to provide the proof of Theorem 3.1 in the following subsection.

4.2 Proof of Theorem 3.1

In this subsection, we present the proof of Theorem 3.1. Before we present a rigorous argument, we first discuss heuristics for the emergence of point attractor. Suppose that the coupling strength K is positive, the initial datum ρ_0 is $\mathcal{C}^1$, and $R_0 > 0$. Then, since R is bounded and monotonically increasing, $\dot{R} \to 0$ as $t \to \infty$ (see Proposition 4.1). It follows from Lemma 4.1(i) that

$$\lim_{t\to\infty} \int_{\mathbb{T}} \sin^2(\theta - \phi(t))\rho(\theta, t)\ d\theta = 0.$$

Thus, the limiting behavior of ρ will be one of the following states: for positive constant $\varepsilon \in (0, \frac{1}{2})$,

$$\delta_{\phi_\infty}, \quad (1-\varepsilon)\delta_{\phi_\infty} + \varepsilon\delta_{(\phi_\infty+\pi)}.$$

We then show that the latter case, i.e., bipolar state is not possible. The proof can be split into two steps:

- Step A: Mass will concentrate asymptotically near at ϕ_∞ and/or $\phi_\infty + \pi$:

$$\lim_{t\to\infty} \int_{\mathbb{T}\setminus I_\delta} \rho(\theta, t)\ d\theta = 0.$$

- Step B: Mass in the interval I_δ^- decays to zero exponentially fast:

$$\lim_{t\to\infty} \int_{I_\delta^-(t)} \rho(\theta, t)\ d\theta = 0,$$

  where the intervals I_δ and $I_\delta^\pm$ are defined in (3.1).

4.2.1 Step A (concentration of mass in the interval I_δ):

In this part, we show that mass will concentrate on the interval I_δ asymptotically. For any $\delta \in (0, \frac{\pi}{2})$, we claim:

$$\lim_{t\to\infty} \int_{\mathbb{T}\setminus I_\delta} \rho(\theta, t)\ d\theta = 0. \tag{4.9}$$

The proof of claim (4.9): It suffices to show that for any $\varepsilon > 0$, there exists a finite time $t_*(\varepsilon) > 0$ such that

$$\int_{\mathbb{T}\setminus I_\delta} \rho(\theta, t)\ d\theta < \varepsilon, \quad t > t_*(\varepsilon).$$

Due to Proposition 4.1, we have $\dot{R} \to 0$ as $t \to \infty$, i.e., there exists a positive time $t_* = t_*(\varepsilon, \delta)$ such that

$$\dot{R}(t) = KR(t) \int_{\mathbb{T}} \sin^2(\theta(t) - \phi(t))\rho(\theta, t)\ d\theta < KR_0(\sin\delta)^2\varepsilon, \quad t > t_*. \tag{4.10}$$

By Lemma 4.1, we have

$$R(t) \geq R_0 \quad t \geq 0. \tag{4.11}$$

Then, it follows from (4.10) and (4.11) that

$$\int_{\mathbb{T}} \sin^2(\theta - \phi(t))\rho(\theta, t)\ d\theta \leq (\sin\delta)^2\varepsilon, \quad t > t_*. \tag{4.12}$$

On the other hand, since

$$|\sin(\theta - \phi(t))| > \sin\delta \quad \forall \theta \in \mathbb{T} \setminus I_\delta,$$

the estimate (4.12) yield

$$\begin{aligned}(\sin\delta)^2 \int_{\mathbb{T}\setminus I_\delta} \rho(\theta, t)\ d\theta &< \int_{\mathbb{T}\setminus I_\delta} \sin^2(\theta - \phi(t))\rho(\theta, t)\ d\theta \\ &\leq \int_{\mathbb{T}} \sin^2(\theta - \phi(t))\rho(\theta, t)\ d\theta \leq (\sin\delta)^2\varepsilon, \quad t \geq t_*.\end{aligned}$$

Thus, we obtain the desired estimate (4.9).

4.2.2 Step B (concentration of mass in the interval I_δ^+):

In this part, we exclude the possibility of bi-polar configuration as an asymptotic profile by showing that no mass concentration occurs in I_δ^- asymptotically, i.e., we claim:

$$\lim_{t\to\infty}\int_{I_\delta^-(t)}\rho(\theta,t)\,d\theta = 0. \tag{4.13}$$

For the proof of (4.13), we use Cauchy-Schwarz inequality to see

$$\begin{aligned}&\int_{I_\delta^-(t)}\rho(\theta,t)\,d\theta\\ &\le\Big(\int_{I_\delta^-(t)}|\rho(\theta,t)|^2\,d\theta\Big)^{\frac{1}{2}}\Big(\int_{I_\delta^-(t)}d\theta\Big)^{\frac{1}{2}}\le\sqrt{2\delta}\Big(\int_{I_\delta^-(t)}|\rho(\theta,t)|^2\,d\theta\Big)^{\frac{1}{2}}.\end{aligned} \tag{4.14}$$

Due to the relation (4.14), it suffices to show that there exist a positive number T_1 such that

$$\int_{I_\delta^-(t)}|\rho(\theta,t)|^2\,d\theta\le e^{-R(0)(\cos\delta)K(t-T_1)}\int_{I_\delta^-(T_1)}|\rho(\theta,T_1)|^2\,d\theta,\qquad\forall\ t>T_1. \tag{4.15}$$

Note that for $\delta\in(0,\frac{\pi}{2})$ and $t>0$,

$$\theta\in I_\delta^-\quad\Longrightarrow\quad\cos(\theta-\phi)<-\cos\delta. \tag{4.16}$$

On the other hand, since $\lim\limits_{t\to\infty}\dot{\phi}=0$, for any $\varepsilon\in(0,K)$, there exist $T_1=T_1(\varepsilon,\delta)>0$ such that

$$|\,\dot{\phi}\,|<\varepsilon R_0\sin\delta,\qquad\forall\ t>T_1. \tag{4.17}$$

We next introduce the Lyapunov functional:

$$\Lambda(t):=\int_{I_\delta^-(t)}|\rho(\theta,t)|^2\,d\theta=\int_{\phi+\pi-\delta}^{\phi+\pi+\delta}|\rho(\theta,t)|^2\,d\theta, \tag{4.18}$$

and we show that it satisfies a Gronwall's inequality:

$$\frac{d\Lambda(t)}{dt} \leq -R(0)\cos\delta K\Lambda(t), \quad t \geq T_1. \tag{4.19}$$

For the estimate (4.19), we use (2.21) and (4.18) to see

$$\begin{aligned}
\frac{d\Lambda}{dt} &= \dot{\phi}\left(\rho^2(\phi+\pi+\delta) - \rho^2(\phi+\pi-\delta)\right) + 2\int_{I_\delta^-} \rho\partial_t\rho \ d\theta \\
&= \dot{\phi}\left(\rho^2(\phi+\pi+\delta) - \rho^2(\phi+\pi-\delta)\right) + 2RK\int_{I_\delta^-} \rho\partial_\theta\Big[\sin(\theta-\phi)\rho\Big] \ d\theta \\
&= \dot{\phi}\left(\rho^2(\phi+\pi+\delta) - \rho^2(\phi+\pi-\delta)\right) - 2RK\sin\delta\rho^2(\phi+\pi+\delta) \\
&- 2RK\sin\delta\rho^2(\phi+\pi-\delta) - RK\int_{I_\delta^-} \sin(\theta-\phi)\partial_\theta(\rho^2) \ d\theta \\
&= \dot{\phi}\rho^2(\phi+\pi+\delta) - \dot{\phi}\rho^2(\phi+\pi-\delta) - RK\sin\delta\rho^2(\phi+\pi+\delta) \\
&- RK\sin\delta\rho^2(\phi+\pi-\delta) + RK\int_{I_\delta^-} \cos(\theta-\phi)\rho^2 \ d\theta \\
&= \Big(\dot{\phi} - RK\sin\delta\Big)\rho^2(\phi+\pi+\delta) - \Big(\dot{\phi} + RK\sin\delta\Big)\rho^2(\phi+\pi-\delta) \\
&+ RK\int_{I_\delta^-} \cos(\theta-\phi)\rho^2 \ d\theta.
\end{aligned} \tag{4.20}$$

On the other hand, it follows from (4.16) and (4.17) that we have

$$\begin{aligned}
&\dot{\phi} - RK\sin\delta \leq (\varepsilon R(0) - RK)\sin\delta \leq (\varepsilon - K)R(0)\sin\delta < 0, \\
&\dot{\phi} + RK\sin\delta > (-\varepsilon R(0) + RK)\sin\delta > (K-\varepsilon)R(0)\sin\delta > 0, \\
&\int_{I_\delta^-} \cos(\theta-\phi)\rho^2 \ d\theta \leq -(\cos\delta)\Lambda(t).
\end{aligned} \tag{4.21}$$

We combine (4.20), (4.21) and the fact $R(t) \geq R_0$ in Proposition 4.1 to obtain (4.19). Thus, we have (4.15). Finally, we combine (4.14) and (4.15) to get

$$\begin{aligned}\int_{I_\delta^-(t)} \rho(\theta,t)\, d\theta &\leq \sqrt{2\delta}\Big(\int_{I_\delta^-(t)} \rho^2(\theta,t)\, d\theta\Big)^{\frac{1}{2}} \\ &\leq \sqrt{2\delta} e^{-\frac{R(0)(\cos\delta)}{2}K(t-T_1)}\Big(\int_{I_\delta^-(T_1)} \rho(\theta,T_1)^2\, d\theta\Big)^{\frac{1}{2}}, \qquad \forall\, t > T_1.\end{aligned}$$

This yields (4.13) and completes the proof of Theorem 3.1.

Remark 4.1. *Note that the result in Theorem 3.1 also implies*

$$\lim_{t\to\infty} R(t) = 1 \quad \textit{and} \quad \lim_{t\to\infty} V_k(\rho) = 0.$$

In the next section, we study existence of positively invariant set for the K-S equation with distributed natural frequencies from well-prepared initial data, in which some significant fraction of mass is confined and it attracts a neighboring mass.

5 Emergence of phase concentration

In this section, we study emergent phenomenon of phase concentration for the K-S equation with distributed natural frequencies, i.e., the non-identical case, from well-prepared initial configurations whose significant portion of mass is already concentrated on the average phase. As we have seen in the previous section, the analysis on the dynamics of global order parameters R and ϕ does play a key role in the proof of the first result in Theorem 3.1. Likewise, we will introduce local order parameters for the sub-ensemble with the same natural frequencies and study the dynamics of these local order parameters. We also discuss a possible asymptotic behavior for the K-S equation with distributed natural frequencies. Finally, we present the proof of our second result Theorem 3.2 on the emergence of arc type attractors from a well aggregated initial datum in a large coupling strength regime. Note that in the next section, such a condition on the initial datum will be removed, but we still are able to show an asymptotic pattern of the mass and the amplitude order parameter which shows asymptotic emergence of complete synchronization, namely, a point cluster, as the coupling strength tends to infinity.

5.1 Local order parameters

For a finite-dimensional Kuramoto model, all oscillators with the same natural frequency will aggregate to the same phase asymptotically. Thus, it is reasonable to consider order parameters for the sub-ensemble of oscillators with the same frequency, which we call *local* order parameters in the sequel. For a fixed $\omega \in \text{supp } g(\cdot)$, let $\varrho(\theta, \omega, t)$ be the conditional probability density function corresponding to the natural frequency ω:

$$f(\theta, \omega, t) = g(\omega)\varrho(\theta, \omega, t) \quad \text{and} \quad \int_{\mathbb{T}} \varrho(\theta, \omega, t)\, d\theta = 1. \tag{5.1}$$

Then, the local order parameters are defined as follows.

Definition 5.1. *Let ϱ be a conditional distribution function introduced in* (5.1)*. Then, for a given $\omega \in$ supp $g(\cdot)$ and $t \geq 0$, the* local *order parame-*

ters R_ω and ϕ_ω are defined by the following relation:

$$R_\omega(t)e^{i\phi_\omega(t)} := \int_{\mathbb{T}} e^{i\theta}\varrho(\theta,\omega,t)\ d\theta. \tag{5.2}$$

Then, the local order parameters satisfy the following estimates.

Lemma 5.1. *Let (R,ϕ) and (R_ω,ϕ_ω) be global and local order parameters defined in* (2.16) *and* (5.2)*, respectively. Then, we have*

$$(i)\ R_\omega = \int_{\mathbb{T}} \varrho(\theta,\omega,t)\cos(\theta-\phi_\omega)\ d\theta, \qquad 0 = \int_{\mathbb{T}} \varrho(\theta,\omega,t)\sin(\theta-\phi_\omega)\ d\theta.$$
$$(ii)\ R = \int_{\mathbb{R}} g(\omega)R_\omega\cos(\phi_\omega-\phi)\ d\omega, \qquad 0 = \int_{\mathbb{R}} g(\omega)R_\omega\sin(\phi_\omega-\phi)\ d\omega.$$

Proof. (i) We divide (5.2) by $e^{i\phi_\omega(t)}$ to get

$$R_\omega(t) = \int_{\mathbb{T}} e^{\mathrm{i}(\theta-\phi_\omega(t))}\varrho(\theta,\omega,t)\ d\theta. \tag{5.3}$$

We now compare the real and imaginary parts of (5.3) to get the desired estimates.

(ii) We use the defining relation (2.16) for R and ϕ to obtain

$$\begin{aligned} Re^{i\phi} &:= \iint_{\mathbb{T}\times\mathbb{R}} f(\theta,\omega,t)e^{i\theta}\ d\theta d\omega = \iint_{\mathbb{T}\times\mathbb{R}} g(\omega)\varrho(\theta,\omega,t)e^{i\theta}\ d\theta d\omega \\ &= \int_{\mathbb{R}} g(\omega)\int_{\mathbb{T}} \varrho(\theta,\omega,t)e^{i\theta}\ d\theta d\omega = \int_{\mathbb{R}} g(\omega)R_\omega e^{i\phi_\omega}\ d\omega. \end{aligned} \tag{5.4}$$

This again yields

$$R = \int_{\mathbb{R}} g(\omega)R_\omega e^{\mathrm{i}(\phi_\omega-\phi)}\ d\omega.$$

We compare real and imaginary parts of the above relation to get the desired estimates. □

We next derive an equation for the conditional probability density function ϱ from the K-S equation (2.15). Recall that f satisfies

$$\begin{aligned}
&\partial_t f + \partial_\theta(v[f]f) = 0, \qquad (\theta, \omega) \in \mathbb{T} \times \mathbb{R}, \quad t > 0, \\
&v[f](\theta, \omega, t) = \omega - K \iint_{\mathbb{T}\times\mathbb{R}} \sin(\theta - \theta_*) f(\theta_*, \omega_*, t)\ d\theta_* d\omega_*.
\end{aligned} \tag{5.5}$$

We now substitute the ansatz (5.1) into the above equation (5.5) to derive the equation for the conditional distribution ϱ:

$$\begin{aligned}
&\partial_t \varrho(\theta, \omega, t) + \omega \partial_\theta \varrho(\theta, \omega, t) \\
&\qquad - K\partial_\theta \Big[\varrho(\theta, \omega, t) \iint_{\mathbb{T}\times\mathbb{R}} \sin(\theta - \theta_*) g(\omega_*) \varrho(\theta_*, \omega_*, t)\ d\theta_* d\omega_* \Big] = 0.
\end{aligned} \tag{5.6}$$

As noticed in (2.19), we have

$$\iint_{\mathbb{T}\times\mathbb{R}} \sin(\theta - \theta_*) g(\omega_*) \varrho(\theta_*, \omega_*, t)\ d\theta_* d\omega_* = R \sin(\theta - \phi).$$

Thus (5.6) can be written as

$$\partial_t \varrho + \partial_\theta \Big[\big(\omega - KR \sin(\theta - \phi) \big) \varrho \Big] = 0. \tag{5.7}$$

This equation can also be obtained directly from the equation (2.20).

Lemma 5.2. *Let* $f = f(\theta, \omega, t)$ *be a solution to* (5.5), *and* (R_ω, ϕ_ω) *and* (R, ϕ) *be local and global order parameters defined by* (2.16) *and* (5.2), *respectively.*

Then, we have

$$(i)\ \dot{R}_\omega = KR \int_{\mathbb{T}} \varrho(\theta, \omega, t) \sin(\theta - \phi_\omega) \sin(\theta - \phi)\ d\theta,$$

$$\dot{\phi}_\omega = \omega - K\frac{R}{R_\omega} \int_{\mathbb{T}} \varrho(\theta, \omega, t) \cos(\theta - \phi_\omega) \sin(\theta - \phi)\ d\theta.$$

$$(ii)\ \dot{R} = -\iint_{\mathbb{T}\times\mathbb{R}} \sin(\theta - \phi)\omega f(\theta, \omega, t)\ d\theta d\omega + KR \int_{\mathbb{T}} \sin^2(\theta - \phi)\rho(\theta, t)\ d\theta.$$

$$\dot{\phi} = \frac{1}{R} \iint_{\mathbb{T}\times\mathbb{R}} \cos(\theta - \phi)\omega f(\theta, \omega, t)\ d\theta d\omega - \frac{K}{2} \int_{\mathbb{T}} \sin\left(2(\theta - \phi)\right) \rho(\theta, t)\ d\theta.$$

Proof. The estimates follow from the differentiation of the defining relations for order parameters and using the K-S equations (5.7) and (5.5) for ρ and f.

(i) We differentiate (5.2) with respect to t and use Lemma 5.1 to obtain

$$\begin{aligned}
\dot{R}_\omega &= \int_{\mathbb{T}} \cos(\theta - \phi_\omega)\partial_t \varrho(\theta, \omega, t)\ d\theta \\
&= -\int_{\mathbb{T}} \cos(\theta - \phi_\omega)\partial_\theta \Big[\varrho(\theta, \omega, t)\big(\omega - KR\sin(\theta - \phi)\big)\Big]\ d\theta \\
&= -\int_{\mathbb{T}} \sin(\theta - \phi_\omega)\varrho(\theta, \omega, t)\big(\omega - KR\sin(\theta - \phi)\big)\ d\theta \\
&= KR\int_{\mathbb{T}} \varrho(\theta, \omega, t)\sin(\theta - \phi_\omega)\sin(\theta - \phi)\ d\theta
\end{aligned}$$

and

$$\begin{aligned}
\dot{\phi}_\omega &= \frac{1}{R_\omega}\int_{\mathbb{T}} \sin(\theta - \phi_\omega)\partial_t \varrho(\theta, \omega, t)\ d\theta \\
&= -\frac{1}{R_\omega}\int_{\mathbb{T}} \sin(\theta - \phi_\omega)\partial_\theta \Big[\varrho(\theta, \omega, t)\big(\omega - KR\sin(\theta - \phi)\big)\Big]\ d\theta \\
&= \frac{1}{R_\omega}\int_{\mathbb{T}} \cos(\theta - \phi_\omega)\varrho(\theta, \omega, t)\big(\omega - KR\sin(\theta - \phi)\big)\ d\theta \\
&= \omega - K\frac{R}{R_\omega}\int_{\mathbb{T}} \varrho(\theta, \omega, t)\cos(\theta - \phi_\omega)\sin(\theta - \phi)\ d\theta.
\end{aligned}$$

(ii) For the estimates on the global order parameters (5.4), we perform similar computation as (i) to obtain desired estimates. $\square$

Remark 5.1. *The dynamics of order parameters* (R_ω, ϕ_ω) *and* (R, ϕ) *coincide with the dynamics* (2.9) *of corresponding order parameters for the Kuramoto model* (2.3)*.*

5.2 Nonexistence of point attractors

In Section 4, we have shown that point attractors can emerge from generic smooth initial data in a positive coupling strength regime for the identical natural frequency case. In this subsection, we will show that emergence of point attractors will not be possible in a general setting. Without loss of generality, we assume that average natural frequencies $\omega_c = \int_{\mathbb{R}} \omega g(\omega) d\omega$ is zero, otherwise, we can consider the rotating frame moving with ω_c.

Suppose that f^∞ is an equilibrium for the K-S equation (5.7), whose conditional probability density function $\varrho^\infty(\theta, \omega, t)$ is in the form $\varrho^\infty(\theta, \omega, t) \equiv \delta_{\phi_\omega}$ for each $\omega \in \operatorname{supp} g$, i.e.,

$$f^\infty(\theta, \omega, t) = g(\omega)\varrho^\infty(\theta, \omega, t) = g(\omega)\delta_{\phi_\omega}. \tag{5.8}$$

We call such f^∞ as a *locally synchronized state*, i.e., a locally synchronized state f^∞ is a complete phase synchronization for a sub-ensemble with the same given frequency ω. A complete phase synchronization is obviously of such a form. To distinguish these locally synchronized states, we use the notation (R^∞, ϕ^∞)

and $(R^\infty_\omega, \phi^\infty_\omega)$ for their global and local order parameters, respectively.

Note that for a locally synchronized state f^∞ in (5.8), it follows from Lemma 5.1 that for each $\omega \in \operatorname{supp} g(\omega)$,

$$R^\infty_\omega = \int_{\mathbb{T}} \delta_{\phi^\infty_\omega} \cos(\theta - \phi^\infty_\omega) d\theta = \cos(\phi^\infty_\omega - \phi^\infty_\omega) = 1. \tag{5.9}$$

This and Lemma 5.2 imply that for all $\omega \in \operatorname{supp}\ g(\omega)$,

$$\begin{aligned}
\dot{\phi}^\infty_\omega &= \omega - K\frac{R^\infty}{R^\infty_\omega}\int_{\mathbb{T}} \varrho^\infty(\theta,\omega,t)\cos(\theta - \phi^\infty_\omega)\sin(\theta - \phi^\infty)\ d\theta \\
&= \omega - KR^\infty \int_{\mathbb{T}} \delta_{\phi^\infty_\omega} \cos(\theta - \phi^\infty_\omega)\sin(\theta - \phi^\infty)\ d\theta \\
&= \omega - KR^\infty \sin(\phi^\infty_\omega - \phi^\infty) \\
&= 0,
\end{aligned} \tag{5.10}$$

where the last line is due to equilibrium state $\dot{\phi}^\infty_\omega = 0$. Thus, for all $\omega \in \operatorname{supp}\ g(\omega)$, we have

$$\sin(\phi^\infty_\omega - \phi^\infty) = \frac{\omega}{KR^\infty}, \quad \text{i.e.,} \quad \phi^\infty_\omega - \phi^\infty = \arcsin\frac{\omega}{KR^\infty}. \tag{5.11}$$

On the other hand, we use Lemma 5.1 and (5.9) to see

$$R^\infty = \int_{\mathbb{R}} g(\omega)R^\infty_\omega \cos(\phi^\infty_\omega - \phi^\infty)\ d\omega = \int_{\mathbb{R}} g(\omega)\sqrt{1 - \left(\frac{\omega}{KR^\infty}\right)^2}\ d\omega. \tag{5.12}$$

Note that the condition in (ii) Lemma 5.1 is automatically satisfied from (5.9) and (5.11):

$$\int_{\mathbb{R}} g(\omega)R^\infty_\omega \sin(\phi^\infty_\omega - \phi^\infty)\ d\omega = \frac{1}{KR^\infty}\int_{\mathbb{R}} \omega g(\omega)\ d\omega = 0,$$

where we used our assumption $\int_{\mathbb{R}} \omega g(\omega)\ d\omega = 0$.

In summary, we have

$$
\begin{aligned}
& f^\infty = g(\omega)\delta_{\phi_\omega^\infty} \quad \text{is an equilibrium} \\
& \qquad \Longleftrightarrow \qquad R^\infty = \int_{\mathbb{R}} g(\omega)\sqrt{1-\Big(\frac{\omega}{KR^\infty}\Big)^2}\, d\omega \quad \text{and} \\
& \quad \sin(\phi_\omega^\infty - \phi^\infty) = \frac{\omega}{KR^\infty}.
\end{aligned}
$$

Proposition 5.1. *Suppose that the coupling strength and g satisfy*

$$K > 0, \quad \int_{\mathbb{R}} g(\omega)\ d\omega = 1, \quad \int_{\mathbb{R}} \omega g(\omega)\ d\omega = 0, \quad g \neq \delta.$$

Then the K-S equation (5.5) *may not have a complete phase synchronization.*

Proof. Suppose that the complete phase synchronization occurs, i.e., there exists an equilibrium f^∞ which corresponds to $R^\infty = 1$. Then, the relation (5.12) yields

$$1 = \int_{\mathbb{R}} g(\omega)\sqrt{1-\Big(\frac{\omega}{K}\Big)^2}\, d\omega. \tag{5.13}$$

However, there exist g and K such that the above relation does not hold. For example, we set

$$g(\omega) = \frac{1}{2}\mathbb{1}_{[-1,1]}, \qquad K = 1.$$

then, the L.H.S. of (5.13) satisfies

$$\int_{\mathbb{R}} g(\omega)\sqrt{1-\Big(\frac{\omega}{K}\Big)^2}\, d\omega = \int_0^1 \sqrt{1-\omega^2}\ d\omega = \arctan\omega\Big|_{\omega=0}^{\omega=1} = \frac{\pi}{4} \neq 1.$$

This is contradictory to the relation (5.13). This shows that the complete phase synchronization may not occur. $\square$

It follows from (5.9), (5.10) and (5.12) that for such a locally synchronized state we have

$$\dot{R}^{\infty} = 0, \qquad \dot{\phi}^{\infty} = 0 \qquad \text{and} \qquad R^{\infty} \to 1 \quad \text{as } K \to \infty.$$

Thus, we can expect that for a large K, equilibrium states are close to a complete phase synchronization. In the following proposition, we give a more quantified version of this.

Proposition 5.2. *Suppose that the probability density function $g = g(\omega)$ satisfies*

$$[-m, m] \subset \operatorname{supp}\, g(\omega) \subset [-M, M], \qquad \int_{\mathbb{R}} \omega g(\omega)\, d\omega = 0, \tag{5.14}$$

and let f^{∞} be an equilibrium to (2.15)*. Then, we have*

$$R^{\infty} \geq \sqrt{1 - \left(\frac{M}{KR^{\infty}}\right)^2}, \qquad R^{\infty} \geq m\Big(\min_{\omega \in [-m,m]} g(\omega)\Big).$$

Proof. (i) It follows from (5.12) that we have

$$\begin{aligned} R^{\infty} &= \int_{\operatorname{supp} g(\omega)} g(\omega) \sqrt{1 - \left(\frac{\omega}{KR^{\infty}}\right)^2}\, d\omega \\ &\geq \sqrt{1 - \left(\frac{M}{KR^{\infty}}\right)^2} \int_{\operatorname{supp} g(\omega)} g(\omega)\, d\omega = \sqrt{1 - \left(\frac{M}{KR^{\infty}}\right)^2}. \end{aligned}$$

(ii) For $K \geq m$, we use (5.12) and (5.14) to obtain

$$\begin{aligned}
R^\infty &\geq \int_{-m}^{m} g(\omega)\sqrt{1-\Big(\frac{\omega}{KR^\infty}\Big)^2}\,d\omega \\
&\geq \Big(\min_{\omega\in[-m,m]} g(\omega)\Big)\int_{-m}^{m}\sqrt{1-\Big(\frac{\omega}{KR^\infty}\Big)^2}\,d\omega \\
&= \Big(\min_{\omega\in[-m,m]} g(\omega)\Big)\Big(m\sqrt{1-\Big(\frac{m}{KR^\infty}\Big)^2}+KR^\infty\arcsin\frac{m}{KR^\infty}\Big) \qquad (5.15)\\
&\geq \Big(\min_{\omega\in[-m,m]} g(\omega)\Big)KR^\infty\arcsin\frac{m}{KR^\infty} \\
&\geq m\Big(\min_{\omega\in[-m,m]} g(\omega)\Big).
\end{aligned}$$

$\square$

Remark 5.2. *For $g(\omega) = \frac{1}{2l}\mathbb{1}_{[-l,l]}$, if we choose*

$$m = M = l,$$

then we have $R^\infty \geq \frac{1}{2}$.

5.3 Proof of Theorem 3.2

As noted in Proposition 5.1, a complete phase synchronization may not occur for the distributed natural frequencies and a complete phase synchronization can be regraded as a concentration phenomenon where full mass concentrates at a single point. Thus, it is still interesting to see

> Under what conditions on parameters and initial data, when does a concentration around the average phase emerge?

This question will be addressed in the sequel.

For $t \geq 0$, we consider the following time-dependent interval L_γ^+ (see Figure 5.2) and mass on $S \subset \mathbb{T}$:

$$L_\gamma^+(t) := \big(\phi(t) - \frac{\pi}{2} + \gamma, \phi(t) + \frac{\pi}{2} - \gamma\big), \qquad \mathcal{M}(S) := \int_{\mathbb{R}} \int_S f(\theta, \omega, t)\ d\theta d\omega,$$

where the constant γ is to be determined later, and we assume that $g = g(\omega)$ is compactedly supported and

$$\text{supp } g(\omega) \subset (-M, M).$$

Note that the length of the time-dependent interval $L_\gamma^+(t)$ equals to $\pi - 2\gamma$ and $L_\gamma^+ = I^+_{\frac{\pi}{2}-\gamma}$.

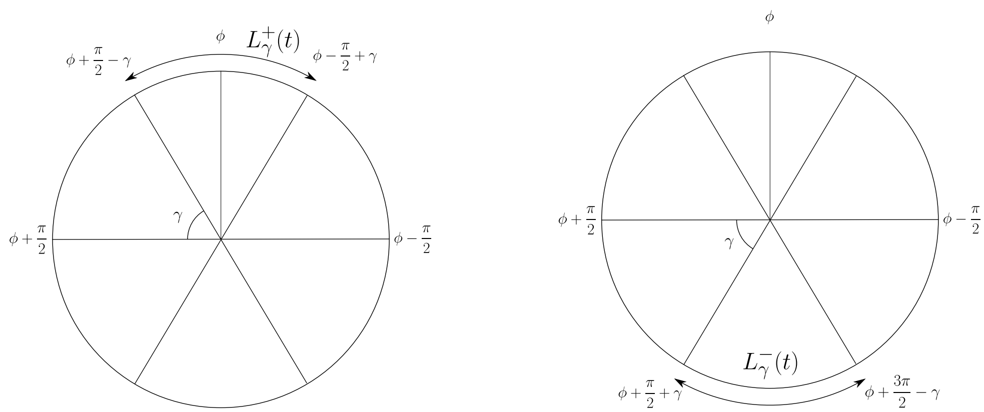
Then, under an appropriate assumption on the initial configuration, we will show the following two properties: for any solution $f = f(\theta, \omega, t)$ to (2.15),

$$\frac{d}{dt}\mathcal{M}(L^+_{\gamma(t)}) \geq 0 \quad \text{and} \quad \lim_{t\to\infty} \int_{L_\gamma^+(t)} |f(\cdot, \omega, t)|^2\ d\theta = \infty, \quad \text{for each } \omega \in \mathbb{R}, \tag{5.16}$$

5.3.1 Verification of the first estimate in (5.16)

Before we present the proof of Theorem 3.2, we first establish several lemmata in the sequel.

We first study the bounds of $\dot{\phi}$ and R.



(a) $L_\gamma^+(t)$ (b) $L_\gamma^-(t)$

Figure 5.2: Geometric descriptions of $L_\gamma^+(t)$ and $L_\gamma^-(t)$

Lemma 5.3. *Let f be a solution to* (2.15)*. Then, the order parameters R and ϕ satisfy*

(i) $|\dot{\phi}| \le \dfrac{M}{R} + K(1-R)$.

(ii) $$\max\{0, (1+\sin\gamma)\mathcal{M}(L_\gamma^+) - 1\} \le R \le \min\{1, (1-\sin\gamma)\mathcal{M}\big(L_\gamma^+\big) + \sin\gamma\}. \tag{5.17}$$

Proof. (i) We use (5.1) and Lemma 5.2 to obtain

$$\begin{aligned} \dot{\phi} &= \frac{1}{R}\int_{\mathbb{R}} \omega g(\omega) \int_{\mathbb{T}} \varrho(\theta,\omega,t)\cos(\theta-\phi)\ d\theta d\omega \\ &\quad - \frac{1}{2}\int_{\mathbb{R}} K g(\omega)\int_{\mathbb{T}} \varrho(\theta,\omega,t)\sin 2(\theta-\phi)\ d\theta d\omega \\ &=: \mathcal{I}_{21} + \mathcal{I}_{22}. \end{aligned} \tag{5.18}$$

- (Estimate of $\mathcal{I}_{21}$): We use the fact supp $g(\omega) \subset [-M, M]$ to see

$$|\mathcal{I}_{21}| \le \frac{1}{R}\int_{\mathbb{R}} |\omega| g(\omega)\int_{\mathbb{T}} \varrho(\theta,\omega,t)\ d\theta d\omega \le \frac{1}{R}\int_{-M}^{M} |\omega| g(\omega)\ d\omega \le \frac{M}{R}. \tag{5.19}$$

• (Estimate of $\mathcal{J}_{22}$): We use the same argument as in the proof of Proposition 4.1 to get

$$|\mathcal{J}_{22}| \le K(1-R). \tag{5.20}$$

Finally, in (5.18), we combine (5.19) and (5.20) to obtain the desired estimate:

$$|\dot{\phi}| \le \frac{M}{R} + K(1-R).$$

(ii) For the lower bound estimate, we use a defining relation (2.16) for R to obtain

$$\begin{aligned}
R(t) &= \iint_{\mathbb{T}\times\mathbb{R}} f(\theta,\omega,t)\cos(\theta-\phi)\ d\theta d\omega \\
&= \int_{\mathbb{R}}\int_{L_\gamma^+(t)} f(\theta,\omega,t)\cos(\theta-\phi)\ d\theta d\omega \\
&+ \int_{\mathbb{R}}\int_{\mathbb{T}\setminus L_\gamma^+(t)} f(\theta,\omega,t)\cos(\theta-\phi)\ d\theta d\omega \\
&\ge \sin\gamma \int_{\mathbb{R}}\int_{L_\gamma^+(t)} f(\theta,\omega,t)\ d\theta d\omega - \int_{\mathbb{R}}\int_{\mathbb{T}\setminus L_\gamma^+(t)} f(\theta,\omega,t)\ d\theta d\omega \\
&= \sin\gamma \mathcal{M}(L_\gamma^+(t)) - \Big(1-\mathcal{M}(L_\gamma^+(t))\Big) \\
&= (1+\sin\gamma)\mathcal{M}(L_\gamma^+(t)) - 1.
\end{aligned}$$

On the other hand, for the upper bound estimate, we use

$$R \le \iint_{L_\gamma^+\times\mathbb{R}} f\ d\theta d\omega + \sin\gamma \iint_{(\mathbb{T}\setminus L_\gamma^+)\times\mathbb{R}} f\ d\theta d\omega \le (1-\sin\gamma)\mathcal{M}(L_\gamma^+) + \sin\gamma$$

to obtain the desired upper bound for R. □

We now find appropriate constants ε_0 and γ_0 for the initial condition.

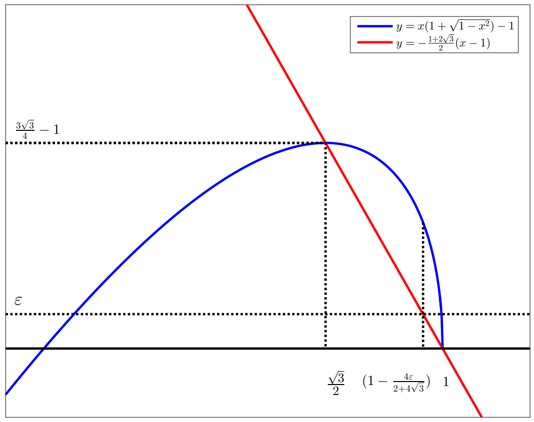


Figure 5.3: $x(1+\sqrt{1-x^2}) - 1 > -\frac{1+2\sqrt{3}}{2}(x-1)$

Lemma 5.4. *Suppose ε_0 and γ_0 are positive constants satisfying*

$$0 < \varepsilon_0 < \frac{3\sqrt{3}}{4} - 1, \qquad \frac{\pi}{3} \leq \gamma_0 < \arcsin\left(1 - \frac{2\varepsilon_0}{2\sqrt{3}+1}\right). \tag{5.21}$$

Then, we have

$$\frac{1+\varepsilon_0}{1+\sin\gamma_0} < \mathcal{M}_*(\varepsilon_0, \gamma_0) < 1, \tag{5.22}$$

where $\mathcal{M}_*(\varepsilon_0, \gamma_0)$ is a positive constant defined by (3.2)*.*

Proof. (i) (First inequality): Since $1 + \varepsilon_0 > \varepsilon_0(1+\cos\gamma_0)$, we have

$$2 + \varepsilon_0 + \cos\gamma_0 = 1 + \cos\gamma_0 + 1 + \varepsilon_0 > (1+\varepsilon_0)(1+\cos\gamma_0)$$

This yields the first inequality.

(ii) (Second inequality): For $\frac{\sqrt{3}}{2} < x < 1$, we have the following inequality (see Figure 5.3):

$$x(1+\sqrt{1-x^2}) - 1 > -\frac{1+2\sqrt{3}}{2}(x-1).$$

Thus, we have

$$x(1+\sqrt{1-x^2})-1>\varepsilon_0, \quad \text{for} \quad \frac{\sqrt{3}}{2}<x<1-\frac{2\varepsilon_0}{1+2\sqrt{3}}. \tag{5.23}$$

On the other hand, by the assumption (5.21),

$$\frac{\sqrt{3}}{2}<\sin\gamma_0<1-\frac{2\varepsilon_0}{2\sqrt{3}+1}$$

We use (5.23) to obtain

$$\sin\gamma_0(1+\cos\gamma_0)>1+\varepsilon_0, \quad \text{equivalently} \quad (1+\sin\gamma_0)(1+\cos\gamma_0)>2+\varepsilon_0+\cos\gamma_0,$$

which implies (5.22) for $\gamma_0>\frac{\pi}{3}$. The case $\gamma_0=\frac{\pi}{3}$ follows from the fact that

$$\mathcal{M}_*(\varepsilon_0,\gamma_0)=1,$$

when $\gamma_0=\frac{\pi}{3}$ and $\varepsilon_0=\frac{3\sqrt{3}}{4}-1$. $\square$

We now see how the previoiusly chosen ε_0 and γ_0 are used to give an appropriate initial configuration.

Lemma 5.5. *Suppose that the initial datum f_0 satisfies*

$$(i)\ f_0(\theta,\omega)=0 \quad \text{in} \quad \mathbb{T}\times(\mathbb{R}\backslash[-M,M]).$$
$$(ii)\ \inf_{\omega\in supp\ g}\int_{L^+_{\gamma_0}(0)} f_0(\theta,\omega)\ d\theta \ge \mathcal{M}_*(\varepsilon_0,\gamma_0),$$

where ε_0 and γ_0 are positive constants as in Lemma 5.4. Then we have

$$\sup_{\omega\in supp\ g}|\phi(0)-\phi_\omega(0)|<\frac{\pi}{2}-\frac{1+\varepsilon_0}{1+\cos\gamma_0}.$$

Proof. By definition of the local order parameter (5.2) and (3.2), we have

$$
\begin{aligned}
&R_\omega(0)\cos(\phi_\omega(0)-\phi(0)) \\
&\quad = \int_{L^+_{\gamma_0}(0)} \cos(\theta-\phi(0)) f_0(\theta,\omega)\ d\theta + \int_{\mathbb{T}\setminus L^+_{\gamma_0}(0)} \cos(\theta-\phi(0)) f_0(\theta,\omega)\ d\theta \\
&\quad \geq \cos(\frac{\pi}{2}-\gamma_0) \int_{L^+_{\gamma_0}(0)} f_0(\theta,\omega)\ d\theta - \int_{\mathbb{T}\setminus L^+_{\gamma_0}(0)} f_0(\theta,\omega)\ d\theta \\
&\quad \geq \sin\gamma_0 \mathcal{M}_*(\varepsilon_0,\gamma_0) - \Big(1-\mathcal{M}_*(\varepsilon_0,\gamma_0)\Big) \\
&\quad = \frac{2+\varepsilon_0+\cos\gamma_0}{1+\cos\gamma_0} - 1 = \frac{1+\varepsilon_0}{1+\cos\gamma_0} > 0.
\end{aligned}
$$

Since $R_\omega(0) \leq 1$,

$$
\cos(\phi_\omega - \phi) > \frac{1+\varepsilon_0}{1+\cos\gamma_0} > \sin\left(\frac{1+\varepsilon_0}{1+\cos\gamma_0}\right) = \cos\left(\frac{\pi}{2} - \frac{1+\varepsilon_0}{1+\cos\gamma_0}\right).
$$

Therefore, we have

$$
|\phi(0)-\phi_\omega(0)| < \frac{\pi}{2} - \frac{1+\varepsilon_0}{1+\cos\gamma_0}.
$$

□

We are now ready to prove the first part of Theorem 3.2. Let γ_0 and ε_0 be positive constants satisfying the relations (5.21), and suppose that $K, g = g(\omega)$ and the initial datum satisfy

$$
\begin{aligned}
&(i)\ \text{supp}\ g(\omega) \subset [-M, M], \qquad K > \frac{M}{\varepsilon_0}\left(1+\frac{1}{\varepsilon_0}\right), \\
&(ii)\ ||f_0||_{L^\infty} < \infty, \qquad \mathcal{M}(L^+_{\gamma_0}(0)) \geq \mathcal{M}_*(\varepsilon_0,\gamma_0).
\end{aligned}
\tag{5.24}
$$

Then, for any classical solution f to (2.15) we will show that

$$
\frac{d}{dt}\mathcal{M}(L^+_{\gamma_0}(t)) \geq 0.
$$

Since the proof is rather lengthy, we split it into several steps. We first note that

$$f(\theta, \omega, t) = 0, \quad \omega \notin [-M, M].$$

• Step A: When $R(t) > 0$, we first establish

$$\begin{aligned}&\frac{d}{dt}\mathcal{M}(L^+_{\gamma_0}(t)) \\ &\qquad \geq \Big[K\big(R(1+\cos\gamma_0) - 1\big) - M\left(1 + \frac{1}{R}\right)\Big]\int_{-M}^{M}\big|B_{-,\omega}(t) + B_{+,\omega}(t)\big|\,d\omega,\end{aligned} \tag{5.25}$$

where $B_{-,\omega}$ and $B_{+,\omega}$ denote the boundary values:

$$B_{-,\omega}(t) := f(\phi(t) - \frac{\pi}{2} + \gamma_0, \omega, t) \quad \text{and} \quad B_{+,\omega}(t) := f(\phi(t) + \frac{\pi}{2} - \gamma_0, \omega, t). \tag{5.26}$$

For the estimate (5.25), we use straightforward calculation to see

$$
\begin{aligned}
&\frac{d}{dt}\mathcal{M}(L^{+}_{\gamma_0}(t))\\
&\quad = \frac{d}{dt}\int_{-M}^{M}\int_{\phi(t)-\frac{\pi}{2}+\gamma_0}^{\phi(t)+\frac{\pi}{2}-\gamma_0} f(\theta,\omega,t)\ d\theta d\omega\\
&\quad = \dot{\phi}(t)\int_{-M}^{M}\big[f(\phi+\frac{\pi}{2}-\gamma_0,t)-f(\phi-\frac{\pi}{2}+\gamma_0,t)\big]\ d\omega\\
&+\int_{-M}^{M}\int_{L_{\gamma_0}(t)}\partial_t f\ \ d\theta d\omega\\
&\quad = \dot{\phi}(t)\int\big[B_{+,\omega}(t)-B_{-,\omega}(t)\big]\ d\omega\\
&\quad -\int_{-M}^{M}\int_{L^{+}_{\gamma_0}(t)}\partial_\theta\Big[f(\theta,\omega,t)\big(\omega-KR\sin(\theta-\phi)\big)\Big]\ d\theta d\omega\\
&\quad = \dot{\phi}(t)\int_{-M}^{M}\big[B_{+,\omega}(t)-B_{-,\omega}(t)\big]\ d\omega\\
&\quad +\int_{-M}^{M}\big[-B_{+,\omega}(t)\big(\omega-KR\sin(\frac{\pi}{2}-\gamma_0)\big)\\
&+B_{-,\omega}(t)\big(\omega-KR\sin(-\frac{\pi}{2}+\gamma_0)\big)\big]\ d\omega
\end{aligned}
\tag{5.27}
$$

By rearranging the terms, we have

$$
\begin{aligned}
&\frac{d}{dt}\mathcal{M}(L^{+}_{\gamma_0}(t))\\
&\quad = \dot{\phi}(t)\int_{-M}^{M}\big[B_{+,\omega}(t)-B_{-,\omega}(t)\big]\ d\omega\\
&\quad +\int_{-M}^{M}\big[-B_{+,\omega}(t)\big(\omega-KR\cos\gamma_0\big)+B_{-,\omega}(t)\big(\omega+KR\cos\gamma_0\big)\big]\ d\omega\\
&\quad = KR\cos\gamma_0\int_{-M}^{M}\big[B_{-,\omega}(t)+B_{+,\omega}(t)\big]\ d\omega\\
&+\int_{-M}^{M}\big(\dot{\phi}(t)-\omega\big)\big[B_{+,\omega}(t)-B_{-,\omega}(t)\big]\ d\omega.
\end{aligned}
$$

In (5.27), we use (5.17) in Lemma 5.3 to obtain (5.25):

$$
\begin{aligned}
&\frac{d}{dt}\mathcal{M}(L^+_{\gamma_0}(t)) \\
&\quad \geq \int_{-M}^{M} \big(KR\cos\gamma_0 - |\dot{\phi}(t)-\omega|\big)\big|B_{-,\omega}(t)+B_{+,\omega}(t)\big|\, d\omega \\
&\quad \geq \left(KR\cos\gamma_0 - M\left(1+\frac{1}{R}\right) - K(1-R)\right)\int_{-M}^{M}\big|B_{-,\omega}(t)+B_{+,\omega}(t)\big|\, d\omega \\
&\quad = \left(K\big(R(1+\cos\gamma_0)-1\big) - M\left(1+\frac{1}{R}\right)\right)\int_{-M}^{M}\big|B_{-,\omega}(t)+B_{+,\omega}(t)\big|\, d\omega.
\end{aligned}
$$

Now observe that

$$
\begin{aligned}
&K\big(R(1+\cos\gamma_0)-1\big) - M\left(1+\frac{1}{R}\right) \\
&\geq K\Big\{\Big((1+\sin\gamma_0)\mathcal{M}\big(L^+_{\gamma_0}(t)\big)-1\Big)(1+\cos\gamma_0)-1\Big\} \\
&- M\Big(1+\frac{1}{(1+\sin\gamma_0)\mathcal{M}\big(L^+_{\gamma_0}(t)\big)-1}\Big) \\
&=: \Delta(t).
\end{aligned}
$$

In the next steps, we will show that $\Delta \geq 0$.

• Step B: Due to the assumption (5.24) (i), we can choose sufficiently small $\eta > 0$ satisfying

$$
K\Big(\varepsilon_0 - \eta(1+\sin\gamma_0)(1+\cos\gamma_0)\Big) - M\Big(1+\frac{1}{\varepsilon_0-\eta(1+\sin\gamma_0)}\Big) > 0. \quad (5.28)
$$

Note that for $\eta = 0$, the relation (5.28) reduces to

$$
K\varepsilon_0 - M\Big(1+\frac{1}{\varepsilon_0}\Big) = \varepsilon_0\Big[K - \frac{M}{\varepsilon_0}\Big(1+\frac{1}{\varepsilon_0}\Big)\Big] > 0 \quad \text{by assumption (i) in (5.24).}
$$

Thus, such η satisfying (5.28) exists.

• Step C: We claim that for $t \in [0,\infty)$,

$$\Delta(t) \geq K\Big(\varepsilon_0 - \eta(1+\sin\gamma_0)(1+\cos\gamma_0)\Big) - M\Big(1 + \frac{1}{\varepsilon_0 - \eta(1+\sin\gamma_0)}\Big)$$
$$\text{and} \quad R(t) \geq \varepsilon_0. \tag{5.29}$$

The proof of claim (5.29): we now define a set $\mathcal{T}_\eta$ and its supremum $T_\eta^* := \sup \mathcal{T}_\eta$:

$$\mathcal{T}_\eta := \Big\{ T \in [0,\infty) : \mathcal{M}\big(L^+_{\gamma_0}(t)\big) > \mathcal{M}_*(\varepsilon_0,\gamma_0) - \eta \quad \text{for all} \quad t \in [0,T) \Big\}.$$

It follows from (5.17), (5.28), and definition of $\mathcal{T}_\eta$ that for $t \in [0, T_\eta^*)$,

$$R(t) \geq (1+\sin\gamma_0)\mathcal{M}\big(L^+_{\gamma_0}(t)\big) - 1 \geq \frac{1+\varepsilon_0}{1+\cos\gamma_0} - \eta(1+\sin\gamma_0) > 0. \tag{5.30}$$

where we used $\eta \ll 1$. By the assumption (ii) in (5.24) and Lipschitz continuity of $\mathcal{M}\big(L^+_{\gamma_0}(t)\big)$ in t (see Appendix B), we can see that the set $\mathcal{T}_\eta$ is nonempty, hence $T_\eta^* \in (0,\infty]$. Suppose that $T_\eta^* < \infty$. Then, we have

$$\lim_{t \to T_\eta^* -} \mathcal{M}\big(L^+_{\gamma_0}(t)\big) = \mathcal{M}_*(\varepsilon_0,\gamma_0) - \eta. \tag{5.31}$$

Again by (5.17), (5.28) and (5.30), for $t \in [0, T_\eta^*)$ we have

$$\Delta \geq K\Big(\varepsilon_0 - \eta(1+\sin\gamma_0)(1+\cos\gamma_0)\Big) - M\Big(1 + \frac{1}{\varepsilon_0 - \eta(1+\sin\gamma_0)}\Big) \geq 0, \tag{5.32}$$

where we used an inequality $1+\varepsilon_0 > \varepsilon_0(1+\cos\gamma_0)$. Thus, the relation (5.32) yields

$$\mathcal{M}(L^+_{\gamma_0}(t)) \geq \mathcal{M}(L^+_{\gamma_0}(0)), \quad t \in [0, T_\eta^*).$$

We let $t \to T_\eta^*$ and use (5.31), assumption (ii) in (5.24) to obtain

$$\mathcal{M}_*(\varepsilon_0,\gamma_0) - \eta = \lim_{t \to T_\eta^*} \mathcal{M}(L^+_{\gamma_0}(t)) \geq \mathcal{M}(L^+_{\gamma_0}(0)) > \mathcal{M}_*(\varepsilon_0,\gamma_0)$$

which is contradictory. Therefore, we have $T^*_\eta = \infty$ and

$$\mathcal{M}\big(L^+_{\gamma_0}(t)\big) > \mathcal{M}_*(\varepsilon_0, \gamma_0) - \eta, \qquad R(t) \geq \varepsilon_0, \quad \forall\ t \in [0, \infty).$$

In fact, the above inequality holds for any $\eta \ll 1$, thus, we have

$$\mathcal{M}\big(L^+_{\gamma_0}(t)\big) \geq \mathcal{M}_*(\varepsilon_0, \gamma_0), \quad \forall\ t \in [0, \infty).$$

We substitute the above relation again into (5.32) to obtain the desired estimate:

$$\Delta(t) \geq K\varepsilon_0 - M\left(1 + \frac{1}{\varepsilon_0}\right), \quad \forall\ \ t \in [0, \infty),$$

which then shows

$$\frac{d}{dt}\mathcal{M}\big(L^+_{\gamma_0}(t)\big) \geq 0, \quad \forall\ t \in [0, \infty).$$

This completes the proof.

5.3.2 The second part of the proof of Theorem 3.2

In this part, we control the L^2-integral of f_ω on the arc L_{γ_0} and show that concentration of mass occurs on $L_{\gamma_0}(t)$ as time goes to infinity, when the coupling strength is large enough, i.e.,

$$\lim_{t\to\infty} \int_{L^+_{\gamma_0}(t)} |f(\theta, \omega, t)|^2\ d\theta = \infty, \quad \text{for each } \omega \in \text{supp } g(\omega).$$

More precisely, under the same assumptions as in the previous part, we have

$$\int_{L^+_{\gamma_0}(t)} |f(\theta, \omega, t)|^2\ d\theta \geq \int_{L^+_{\gamma_0}(t)} |f_0(\theta, \omega, 0)|^2\ d\theta\ e^{(K\varepsilon_0 \sin\gamma_0)t}, \quad \forall\ \omega \in \text{supp } g(\omega).$$

First, for each $t \geq 0$ and each ω in $[-M, M]$, we define

$$\Gamma^+_{\gamma_0,\omega}(t) := \int_{L^+_{\gamma_0}(t)} |f(\theta, \omega, t)|^2\ d\theta = \int_{\phi(t)-\frac{\pi}{2}+\gamma_0}^{\phi(t)+\frac{\pi}{2}-\gamma_0} |f(\theta, \omega, t)|^2\ d\theta.$$

Then, by direct computation,

$$\begin{aligned}\frac{d}{dt}\Gamma^+_{\gamma_0,\omega}(t) &= \dot\phi(t)\big(B_{+,\omega}(t)\big)^2 - \big(B_{-,\omega}(t)\big)^2 + 2\int_{L^+_{\gamma_0}(t)} f\partial_t f\ d\theta \\ &=: \dot\phi(t)\Big[\big(B_{+,\omega}(t)\big)^2 - \big(B_{-,\omega}(t)\big)^2\Big] + \mathcal{I}_3,\end{aligned} \tag{5.33}$$

where $B_{\pm,\omega}$ is the boundary value defined in (5.26).

For the estimate of $\mathcal{I}_3$, we use (5.7) to obtain

$$\begin{aligned}\mathcal{I}_3(t) &= -2\int_{L^+_{\gamma_0}(t)} f\partial_\theta\Big[f\big(\omega - KR\sin(\theta-\phi)\big)\Big]\ d\theta \\ &= -2\int_{L^+_{\gamma_0}(t)} (f\partial_\theta f)\big(\omega - KR\sin(\theta-\phi)\big) - f^2KR\cos(\theta-\phi)\ d\theta \\ &= -\int_{L^+_{\gamma_0}(t)} (\partial_\theta f^2)\big(\omega - KR\sin(\theta-\phi)\big)\ d\theta + 2KR\int_{L^+_{\gamma_0}(t)} f^2\cos(\theta-\phi)\ d\theta \\ &=: \mathcal{I}_{31}(t) + \mathcal{I}_{32}(t).\end{aligned} \tag{5.34}$$

Below, we estimate the terms $\mathcal{I}_{3i},\ i = 1,2$ separately.

• (Estimate on $\mathcal{I}_{31}$): Integration by parts yields

$$\begin{aligned}\mathcal{I}_{31}(t) = &-\Big[\big(B_{+,\omega}(t)\big)^2\big(\omega - KR\sin(\frac{\pi}{2}-\gamma_0)\big) \\ &- \big(B_{-,\omega}(t)\big)^2\big(\omega - KR\sin(-\frac{\pi}{2}+\gamma_0)\big)\Big] \\ &- KR\int_{L^+_{\gamma_0}(t)} \big(f(\theta,\omega,t)\big)^2\cos(\theta-\phi)\ d\theta \\ &=: \mathcal{I}_{311}(t) + \mathcal{I}_{312}(t).\end{aligned} \tag{5.35}$$

$\diamond$ (Estimate on $\mathcal{I}_{311}$): By rearranging terms, we have

$$\mathcal{I}_{311}(t) = -\omega\Big[\big(B_{+,\omega}(t)\big)^2 - \big(B_{-,\omega}(t)\big)^2\Big] + KR\cos\gamma_0\Big[\big(B_{+,\omega}(t)\big)^2 + \big(B_{-,\omega}(t)\big)^2\Big],$$
$$\mathcal{I}_{32}(t) + \mathcal{I}_{312}(t) \geq KR\sin\gamma_0\Gamma^{+}_{\gamma_0,\omega}(t). \tag{5.36}$$

In (5.33), we combine all estimates (5.34), (5.35), (5.36) and use $R \geq \varepsilon_0$ to obtain

$$\begin{aligned}
&\frac{d}{dt}\Gamma^{+}_{\gamma_0,\omega}(t)\\
&\quad\geq (\dot{\phi}(t)-\omega)\Big[\big(B_{+,\omega}(t)\big)^2 - \big(B_{-,\omega}(t)\big)^2\Big]\\
&\quad+ KR\cos\gamma_0\Big[\big(B_{+,\omega}(t)\big)^2 + \big(B_{-,\omega}(t)\big)^2\Big] + KR\sin\gamma_0\Gamma^{+}_{\gamma_0,\omega}(t)\\
&\quad\geq \big(KR\cos\gamma_0 - |\dot{\phi}(t)-\omega|\big)\Big[\big(B_{+,\omega}(t)\big)^2 + \big(B_{-,\omega}(t)\big)^2\Big] + KR\sin\gamma_0\Gamma^{+}_{\gamma_0,\omega}(t)\\
&\quad\geq \underbrace{\big(KR\cos\gamma_0 - |\dot{\phi}(t)-\omega|\big)}_{=:\tilde{\Delta}(t)}\Big[\big(B_{+,\omega}(t)\big)^2 + \big(B_{-,\omega}(t)\big)^2\Big] + K\varepsilon_0\sin\gamma_0\Gamma^{+}_{\gamma_0,\omega}(t).
\end{aligned} \tag{5.37}$$

We next estimate the sign of $\tilde{\Delta}$. It follows from (5.17) and $|\omega| \leq M$ that we have

$$|\dot{\phi}-\omega| \leq |\dot{\phi}| + M \leq M\Big(1+\frac{1}{R}\Big) + K(1-R). \tag{5.38}$$

Then, (5.38) and (5.17)(ii) imply

$$\begin{aligned}
\tilde{\Delta} &\geq K\Big[R(1+\cos\gamma_0)-1\Big] - M\Big(1+\frac{1}{R}\Big) \\
&\geq K\Big(\big((1+\sin\gamma_0)\mathcal{M}\big(L^+_{\gamma_0}(t)\big)-1\big)(1+\cos\gamma_0)-1\Big) \\
&- M\Bigg(1+\frac{1}{(1+\sin\gamma_0)\mathcal{M}\big(L^+_{\gamma_0}(t)\big)-1}\Bigg) \\
&\geq K\Big(\big((1+\sin\gamma_0)\mathcal{M}\big(L^+_{\gamma_0}(0)\big)-1\big)(1+\cos\gamma_0)-1\Big) \\
&- M\Bigg(1+\frac{1}{(1+\sin\gamma_0)\mathcal{M}\big(L^+_{\gamma_0}(0)\big)-1}\Bigg) \\
&\geq K\varepsilon_0 - M\Big(1+\frac{1}{\varepsilon_0}\Big) > 0.
\end{aligned} \tag{5.39}$$

In the last line, we used the same argument as in the proof of Step B and Step C in Theorem 3.2. Finally, we use (5.37) and (5.39) to obtain a Gronwall's inequality:

$$\frac{d}{dt}\Gamma^+_{\gamma_0,\omega}(t) \geq K\sin\gamma_0\varepsilon_0\Gamma^+_{\gamma_0,\omega}(t).$$

This yields the desired estimate. This completes the proof of Theorem 3.2.

6 Lower bounds for the amplitude order parameter

In this section, we prove Theorem 3.3, by establishing the promised asymptotic lower bound on the order parameter R (assuming $R_0 > 0$). A first key step is the existence of a positive lower bound $\underline{R}$ of the order parameter R for the system (2.15) with $R(0) = R_0$. For such a lower bound, we need a large coupling strength K, depending on $\frac{1}{R_0}$, and under this assumption, we will first show that if $R_0 > 0$, then we can guarantee that the mass in the sector $L^+_{\pi/3}(t)$ will remain above a universal value, whenever $\dot{R} \leq 0$. This will

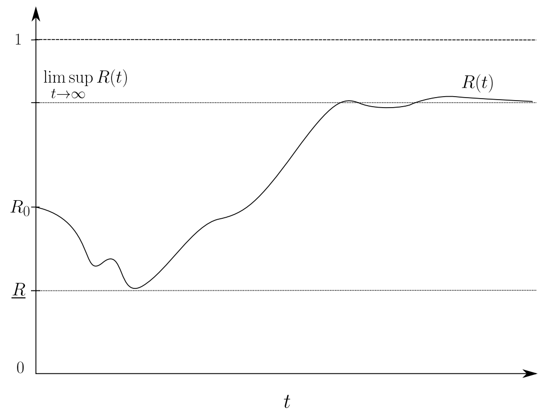


Figure 5.4: Schematic diagram on the dynamics of $R(t)$

enable us to establish the lower bound $\underline{R}$ of R. This is the result of Corollary 6.1, where we also prove that $\dot{R}$ will remain below a small positive constant after some time (see Figure 5.4). This lower bound will induce in Section 6.3. a rough asymptotic lower bound $2/3$ for R; see Proposition 6.2. And we finally improve this rough lower bound to our desired one, namely, R_∞ in Theorem 3.3, which tends to 1 as $K \to \infty$.

Throughout this section, we will assume that the natural frequency density function $g = g(\omega)$ is compactly supported on the interval $[M, M]$. It is important to note that for the results in this section, we do not require the previous assumption $(5.24)(ii)$ given in Section 5 on the initial configuration.

6.1 Several lemmata

In this subsection, for a uniform lower bound of R, we will first present several lemmata. The constants appearing in all computations are not necessarily optimal. Our strategy to find a uniform lower bound is as follows. We will first assume that such a uniform lower bound $\underline{R}$ exists a priori and then later, by the choice of large K, we will remove this a priori assumption and obtain the uniform positive lower bound. We first begin a series of lemmata with the growth estimate for f. We assume that there exists a uniform positive lower bound for R in the following lemmata.

Lemma 6.1. *Let $f = f(\theta, \omega, t)$ be the solution to* (2.15). *Then, we have*

$$\|f(t)\|_{L^\infty(\mathbb{T}\times[-M,M])} \le ||f_0||_{L^\infty} e^{Kt}, \quad t \in [0, \infty).$$

Proof. For $(\theta_0, \omega_0) \in \mathbb{T}\times\mathbb{R}$, we define a forward characteristics $(\theta(t), \omega_0)$ issued from (θ_0, ω_0) at time $t = 0$ as a solution to the following Cauchy problem:

$$\begin{cases} \dot\theta(t) = \omega_0 - KR(t)\sin(\theta(t) - \phi(t)), \quad \dot\omega(t) = 0, \quad t > 0, \\ (\theta(0), \omega(0)) = (\theta_0, \omega_0). \end{cases}$$

Since the right-hand side of the ODE is Lipschitz continuous and sub-linear in (θ, ω), we have a global solution and $\omega(t) = \omega_0$. On the other hand, we use

$$\begin{aligned}(\partial_t f)(\theta, \omega, t) &= -\partial_\theta\big([\omega - KR\sin(\theta - \phi)]f\big) \\ &= -(\omega - KR\sin(\theta - \phi))\partial_\theta f + KR\cos(\theta - \phi)f(\theta, \omega, t).\end{aligned}$$

to see the time-rate of change of f along the characteristics $(\theta(t), \omega_0, t)$:

$$\begin{aligned}&\frac{d}{dt}f(\theta(t), \omega_0, t) \\ &\quad= -[\omega_0 - KR(t)\sin(\theta(t) - \phi(t))]\partial_\theta f(\theta(t), \omega_0, t) \\ &\quad+ KR(t)\cos(\theta(t) - \phi(t))f(\theta(t), \omega_0, t) + \dot\theta(t)\partial_\theta f(\theta(t), \omega_0, t) \\ &\quad= KR(t)\cos(\theta(t) - \phi(t))f(\theta(t), \omega_0, t) \\ &\quad\le Kf(\theta(t), \omega_0, t).\end{aligned} \tag{6.1}$$

This yields

$$f(\theta(t), \omega_0, t) \le e^{Kt} f_0(\theta_0, \omega_0) \le \|f_0\|_{L^\infty(\mathbb{T}\times\mathbb{R})} e^{Kt}, \quad \text{in } [0, T].$$

This yields the desired L^∞-estimate of f. $\square$

Lemma 6.2. *The following assertions hold.*

1. *$R = R(t)$ is Lipschitz continuous in $[0, \infty)$.*

2. *Suppose that the initial density f_0 and the order parameter R satisfy the following conditions:*

$$\|f_0\|_{L^\infty} < \infty, \qquad \inf_{0 \le t \le T} R(t) \ge \underline{R}, \quad \text{for some } T \in (0, \infty] \text{ and } \underline{R} > 0,$$

Then, there exists $\eta' > 0$ such that the functions $R(\cdot), \dot{R}(\cdot)$ and $\mathcal{M}(L^+_\gamma(\cdot))$ are Lipschitz continuous in $[0, T+\eta')$, for any γ in $(-\frac{\pi}{2}, \frac{\pi}{2})$. The Lipschitz constants for $R(\cdot), \dot{R}(\cdot)$ and $(L^+_\gamma(\cdot))$ are given by

$$\begin{aligned}
&|\dot{R}| < M + K,\\
&|\ddot{R}| \le \frac{1}{\underline{R}}\Big[4(K^2 + KM) + 2KM + 2M^2 + 2(M+K)^2\Big],\\
&\left|\frac{d}{dt}\mathcal{M}(L^+_\gamma(t))\right| \le 2M\left[K + \frac{2M}{\underline{R}} + K\Big(1 - \frac{\underline{R}}{2}\Big) + M\right]\|f_0\|_{L^\infty(\mathbb{T}\times\mathbb{R})} e^{K(T+\eta)}.
\end{aligned}$$

Proof. Since the proof is very lengthy, we postpone it to Appendix B. $\square$

In the next Lemma we will show how the values R and $\dot{R}$ can be used to control the mass in $L^+_{\frac{\pi}{3}}$. For $t \geq 0$ and $\gamma \in (\frac{\pi}{3}, \frac{\pi}{2})$, we set

$$
\begin{aligned}
&L^-_\gamma(t) := \big(\phi(t) + \frac{\pi}{2} + \gamma, \phi(t) + \frac{3\pi}{2} - \gamma\big) = I^-_{\frac{\pi}{2}-\gamma},\\
&\tilde{B}_{-,\omega}(t) := f\Big(\phi(t) + \frac{\pi}{2} + \gamma, \omega, t\Big), \quad \tilde{B}_{+,\omega}(t) := f\Big(\phi(t) + \frac{3\pi}{2} - \gamma, \omega, t\Big),\\
&E_1(K, M, \mu, \underline{R}, \gamma) := \frac{\sin\gamma}{\cos^2\gamma}\frac{M}{K\underline{R}} + \frac{1 - \sin\gamma}{2} \quad \text{and}\\
&E_2(K, M, \mu, \underline{R}, \gamma) := 1 - \sin\gamma + (1 + \sin\gamma)\frac{M}{K\underline{R}\cos^2\gamma} + \frac{\sin\gamma}{\cos^2\gamma}\frac{M}{K\underline{R}}.
\end{aligned}
$$

In the subsequent three lemmata, we will study the relationships between $(R(t_0), \dot{R}(t_0))$ and $\mathcal{M}(L^+_{\frac{\pi}{3}}(t_0))$ under the following three situations:

$$
\begin{aligned}
&\text{Case 1}: \ R(t_0) \geq \underline{R}, \qquad \dot{R}(t_0) \leq 0,\\
&\text{Case 2}: \ R(t_0) \geq \underline{R}, \qquad \dot{R}(t_0) < K\mu, \quad K \gg 1, \ \ 0 < \mu \ll 1,\\
&\text{Case 3}: \ \inf_{0 \leq t \leq T} R(t) \geq \underline{R}, \qquad \dot{R}(T) = K\mu.
\end{aligned}
$$

In the sequel, to simplify presentation appearing in the messy computations, we will consider the sector $L^+_{\frac{\pi}{3}}$ and to emphasize γ dependence in E_1 and E_2, we suppress other dependence, i.e. $E_i(\gamma) := E_i(K, M, \mu, \underline{R}, \gamma)$.

Lemma 6.3. *Let $\gamma \in (\frac{\pi}{3}, \frac{\pi}{2})$ and suppose that there exists $t_0 \geq 0$ and $\underline{R} > 0$ such that*

$$R(t_0) \geq \underline{R}, \qquad \dot{R}(t_0) \leq 0. \tag{6.2}$$

Then, $R(t_0)$ is controlled by the mass $\mathcal{M}(L^+_{\frac{\pi}{3}}(t_0))$, and vice verse:

$$2\mathcal{M}(L^+_{\frac{\pi}{3}}(t_0)) - E_2(\gamma) - 1 \leq R(t_0) \leq 2\mathcal{M}(L^+_{\frac{\pi}{3}}(t_0)) + 2E_1(\gamma) - 1. \tag{6.3}$$

Proof. (i) (Proof of the upper bound): For derivation of the second inequality, we first estimate how the mass $\mathcal{M}(L_\gamma^-)$ can be controlled by the mass $\mathcal{M}(L_\gamma^+), K, M, \underline{R}$ and γ. In the sequel, all quantities will be evaluated at $t = t_0$.

• Step A (Controlling the mass $\mathcal{M}(L_\gamma^-)$): By a priori condition (6.2) and Lemma 5.2,

$$\begin{aligned} 0 \geq \dot{R} &= -\iint_{\mathbb{T}\times\mathbb{R}} \sin(\theta-\phi)\omega f(\theta,\omega)\, d\theta d\omega \\ &\quad + KR \iint_{\mathbb{T}\times\mathbb{R}} \sin^2(\theta-\phi) f(\theta,\omega)\, d\theta d\omega \\ &\geq -M + K\underline{R}\cos^2\gamma(1-\mathcal{M}(L_\gamma^+)-\mathcal{M}(L_\gamma^-)), \end{aligned} \tag{6.4}$$

where we used

$$\begin{aligned} &\iint_{\mathbb{T}\times\mathbb{R}} \sin^2(\theta-\phi) f(\theta,\omega)\, d\theta d\omega \\ &\quad \geq \iint_{(\mathbb{T}\setminus(L_\gamma^+\cup L_\gamma^-))\times\mathbb{R}} \sin^2(\theta-\phi) f(\theta,\omega)\, d\theta d\omega \\ &\quad \geq \cos^2\gamma \iint_{(\mathbb{T}\setminus(L_\gamma^+\cup L_\gamma^-))\times\mathbb{R}} f(\theta,\omega)\, d\theta d\omega \\ &\quad = \cos^2\gamma\Big(1-\mathcal{M}(L_\gamma^+)-\mathcal{M}(L_\gamma^-)\Big). \end{aligned}$$

Then the relation (6.4) yields

$$\begin{aligned} &1-\mathcal{M}(L_\gamma^+(t_0))-\mathcal{M}(L_\gamma^-(t_0)) \leq \frac{M}{K\underline{R}\cos^2\gamma}, \quad \text{or} \\ &\mathcal{M}(L_\gamma^-(t_0)) \geq 1-\mathcal{M}(L_\gamma^+(t_0)) - \frac{M}{K\underline{R}\cos^2\gamma}. \end{aligned} \tag{6.5}$$

• Step B (Bounding R by $\mathcal{M}(L_\gamma^+)$): Since $\frac{\pi}{3} < \gamma$, we have

$$L_\gamma^+ \subset L_{\frac{\pi}{3}}^+.$$

Then, we use the above relation, (2.17) and (6.5) to obtain

$$
\begin{aligned}
R &= \iint_{L_\gamma^+\times\mathbb{R}} \langle e^{\mathrm{i}\theta}, e^{\mathrm{i}\phi}\rangle f \ d\theta d\omega + \iint_{(\mathbb{T}\setminus(L_\gamma^+\cup L_\gamma^-))\times\mathbb{R}} \langle e^{\mathrm{i}\theta}, e^{\mathrm{i}\phi}\rangle f \ d\theta d\omega \\
&\quad + \iint_{L_\gamma^-\times\mathbb{R}} \langle e^{\mathrm{i}\theta}, e^{\mathrm{i}\phi}\rangle f \ d\theta d\omega \\
&\le \mathcal{M}(L_\gamma^+) + \sin\gamma\Big(1-\mathcal{M}(L_\gamma^-)-\mathcal{M}(L_\gamma^+)\Big) - \sin\gamma\mathcal{M}(L_\gamma^-) \\
&\le \mathcal{M}(L_\gamma^+) + \frac{\sin\gamma}{\cos^2\gamma}\frac{M}{K\underline{R}} - \sin\gamma\mathcal{M}(L_\gamma^-) \\
&\le \mathcal{M}(L_\gamma^+) + 2\frac{\sin\gamma}{\cos^2\gamma}\frac{M}{K\underline{R}} + \sin\gamma\mathcal{M}(L_\gamma^+) - \sin\gamma \\
&\le 2\mathcal{M}(L_\gamma^+) - 1 + 2\frac{\sin\gamma}{\cos^2\gamma}\frac{M}{K\underline{R}} + (\sin\gamma-1)\mathcal{M}(L_\gamma^+) + 1 - \sin\gamma \\
&\le 2\mathcal{M}(L_{\frac{\pi}{3}}^+) - 1 + 2\frac{\sin\gamma}{\cos^2\gamma}\frac{M}{K\underline{R}} + 1 - \sin\gamma \\
&= 2\Big(\mathcal{M}(L_{\frac{\pi}{3}}^+) - \frac{1}{2} + \frac{\sin\gamma}{\cos^2\gamma}\frac{M}{K\underline{R}} + \frac{1-\sin\gamma}{2}\Big) \\
&= 2\Big(\mathcal{M}(L_{\frac{\pi}{3}}^+) - \frac{1}{2} + E_1\Big).
\end{aligned} \tag{6.6}
$$

This verifies the upper bound.

(ii) (Proof of the lower bound): We use the similar argument to the first part of (6.6) to find

$$
\begin{aligned}
R &= \iint_{L_\gamma^+\times\mathbb{R}} \langle e^{\mathrm{i}\theta}, e^{\mathrm{i}\phi}\rangle f \ d\theta d\omega + \iint_{(\mathbb{T}\setminus(L_\gamma^-\cup L_\gamma^+))\times\mathbb{R}} \langle e^{\mathrm{i}\theta}, e^{\mathrm{i}\phi}\rangle f \ d\theta d\omega \\
&\quad + \iint_{L_\gamma^-\times\mathbb{R}} \langle e^{\mathrm{i}\theta}, e^{\mathrm{i}\phi}\rangle f \ d\theta d\omega \\
&\ge \sin\gamma\mathcal{M}(L_\gamma^+) - \sin\gamma\Big(1-\mathcal{M}(L_\gamma^+)-\mathcal{M}(L_\gamma^-)\Big) - \mathcal{M}(L_\gamma^-).
\end{aligned} \tag{6.7}
$$

On the other hand, we again use the same arguments as in (6.4) to find

$$
\begin{aligned}
0 \geq \dot{R} = &- \iint_{\mathbb{T}\times\mathbb{R}} \sin(\theta-\phi)\omega f(\theta,\omega,t)\ d\theta d\omega \\
&+ KR \iint_{\mathbb{T}\times\mathbb{R}} \sin^2(\theta-\phi)\big)\rho(\theta,t)\ d\theta d\omega \\
&\geq -M + K\underline{R}\cos^2\gamma(1-\mathcal{M}(L_\gamma^+)-\mathcal{M}(L_\gamma^-)).
\end{aligned}
$$

This yields

$$
1-\mathcal{M}(L_\gamma^+)-\mathcal{M}(L_\gamma^-) \leq \frac{M}{K\underline{R}\cos^2\gamma}.
$$

We use (6.7) and the fact that $\mathcal{M}(L_\gamma^+)+\mathcal{M}(L_\gamma^-) \leq 1$ to obtain

$$
\begin{aligned}
R &\geq \sin\gamma\mathcal{M}(L_\gamma^+) - \frac{\sin\gamma}{\cos^2\gamma}\frac{M}{K\underline{R}} - \mathcal{M}(L_\gamma^-) \\
&\geq \sin\gamma\mathcal{M}(L_\gamma^+) - \frac{\sin\gamma}{\cos^2\gamma}\frac{M}{K\underline{R}} - 1 + \mathcal{M}(L_\gamma^+) \\
&= (1+\sin\gamma)\mathcal{M}(L_\gamma^+) - 1 - \frac{\sin\gamma}{\cos^2\gamma}\frac{M}{K\underline{R}}.
\end{aligned}
$$

Similarly, Lemma 5.2 and the fact that $\dot{R} \leq 0$ imply

$$
\mathcal{M}(L_{\frac{\pi}{3}}^+) - \frac{M}{K\underline{R}\cos^2\gamma} \leq \mathcal{M}(L_\gamma^+).
$$

Hence,

$$
R \geq (1+\sin\gamma)\mathcal{M}(L_{\frac{\pi}{3}}^+) - 1 - (1+\sin\gamma)\frac{M}{K\underline{R}\cos^2\gamma} - \frac{\sin\gamma}{\cos^2\gamma}\frac{M}{K\underline{R}}.
$$

Thus, we again use the fact that $\mathcal{M}(L_{\frac{\pi}{3}}^+) \leq 1$ to get

$$
R \geq 2\mathcal{M}(L_{\frac{\pi}{3}}^+) - 1 - (1-\sin\gamma) - (1+\sin\gamma)\frac{M}{K\underline{R}\cos^2\gamma} - \frac{\sin\gamma}{\cos^2\gamma}\frac{M}{K\underline{R}}.
$$

This yields the desired result. $\square$

Remark 6.1. *Note that the estimate* (6.3) *can be rewritten as follows.*

$$\frac{R(t_0)+1}{2} - E_1(\gamma) \le \mathcal{M}(L^+_{\frac{\pi}{3}}(t_0)) \le \frac{R(t_0)+1}{2} + \frac{E_2(\gamma)}{2}.$$

We next show that there exists a positive constant μ, such that when $\dot{R}$ is below $K\mu$, the mass in $L^+_{\frac{\pi}{3}}$ is nondecreasing.

Lemma 6.4. *Let $K, \underline{R}$ and μ satisfy the relation:*

$$\frac{1}{2} > \frac{M}{K\underline{R}} + \frac{M}{K\underline{R}^2} + \frac{1}{\underline{R}\sqrt{\underline{R}}}\sqrt{\frac{M}{K} + \mu}, \tag{6.8}$$

that is, K is sufficiently large and μ is sufficiently small relative to $\underline{R}$. Suppose that at $t = t_0$, the order parameter R satisfies

$$R(t_0) \ge \underline{R}, \qquad \dot{R}(t_0) < K\mu.$$

Then, we have

$$\frac{d}{dt}\Big|_{t=t_0} \mathcal{M}(L^+_{\frac{\pi}{3}}(t)) \ge 0.$$

Proof. In the sequel, for notational simplicity, we will assume that all the time dependent expression are evaluated at $t = t_0$. By (5.27), we have

$$\begin{aligned} \frac{d}{dt}\mathcal{M}(L^+_{\frac{\pi}{3}}(t)) &= KR\cos\frac{\pi}{3}\int_{-M}^{M}\big[B_{-,\omega}(t) + B_{+,\omega}(t)\big]\,d\omega \\ &\quad + \int_{-M}^{M}\big(\dot{\phi}(t) - \omega\big)\big[B_{+,\omega}(t) - B_{-,\omega}(t)\big]\,d\omega \\ &\ge \left(\frac{K\underline{R}}{2} - M - |\dot{\phi}|\right)\int_{-M}^{M}\big[B_{+,\omega}(t) + B_{-,\omega}(t)\big]\,d\omega. \end{aligned} \tag{6.9}$$

Thus, we need to show

$$\frac{K\underline{R}}{2} - M - |\dot{\phi}| \geq 0. \tag{6.10}$$

To check (6.10), we use Lemma 5.2 to find

$$\begin{aligned}
\dot{\phi} &= \frac{1}{R} \iint_{\mathbb{T}\times\mathbb{R}} \cos(\theta - \phi)\omega f(\theta, \omega, t)\ d\theta d\omega \\
&\quad - \frac{K}{2} \iint_{\mathbb{T}\times\mathbb{R}} \sin\left(2(\theta - \phi)\right) f(\theta, \omega, t)\ d\theta d\omega, \\
\dot{R} &= - \iint_{\mathbb{T}\times\mathbb{R}} \sin(\theta - \phi)\omega f(\theta, \omega, t)\ d\theta d\omega \\
&\quad + KR \iint_{\mathbb{T}\times\mathbb{R}} \sin^2(\theta - \phi)\Big) f(\theta, \omega, t)\ d\theta d\omega.
\end{aligned} \tag{6.11}$$

Then, we use (6.11), $f = 0$ for $|\omega| > M$ and Cauchy-Schwarz inequality to get

$$\begin{aligned}
|\dot{\phi}| &< \frac{M}{R} + K \iint_{\mathbb{T}\times\mathbb{R}} |\sin(\theta - \phi)| f(\theta, \omega, t)\ d\theta d\omega. \\
&< \frac{M}{R} + \sqrt{\frac{K}{R}} \sqrt{KR \iint_{\mathbb{T}\times\mathbb{R}} \sin^2(\theta - \phi) f(\theta, \omega, t)\ d\theta d\omega} \\
&< \frac{M}{R} + \sqrt{\frac{K}{R}} \sqrt{M + \dot{R}}.
\end{aligned} \tag{6.12}$$

Note that the conditions (6.8) yield

$$\frac{1}{2} > \frac{M}{K\underline{R}} + \frac{M}{K\underline{R}^2} + \frac{1}{\underline{R}\sqrt{\underline{R}}} \sqrt{\frac{M}{K} + \mu} > \frac{M}{K\underline{R}} + \frac{M}{K\underline{R}^2} + \frac{1}{\underline{R}\sqrt{\underline{R}}} \sqrt{\frac{M}{K} + \frac{\dot{R}}{K}}.$$

By multiplying $K\underline{R}$, we have

$$\frac{K\underline{R}}{2} - M > \frac{M}{\underline{R}} + \sqrt{\frac{K}{\underline{R}}} \sqrt{M + \dot{R}}. \tag{6.13}$$

We now combine (6.12) and (6.13) to get

$$|\dot{\phi}(t_0)| < \frac{K\underline{R}}{2} - M.$$

This and (6.9) implies the desired estimate. $\square$

In the following Lemma, under the assumption that $\dot{R}(t) \geq K\mu$ for some t, we quantify the increase of R, at some later time.

Lemma 6.5. *Suppose that f_0, R and K satisfy*

$$\|f_0\|_{L^\infty(\mathbb{T}\times\mathbb{R})} < \infty, \qquad \inf_{0\leq t\leq T} R(t) \geq \underline{R}, \qquad \dot{R}(T) = K\mu,$$
$$K^2\mu > \frac{M^2}{2\underline{R}^2} - \frac{3M^2}{4\underline{R}}$$

for some $T \geq 0$ and some positive constants $\underline{R}$, and μ. Then there exist positive constants $d := d(K, M, \underline{R}, \mu)$ and $E_3 := E_3(K, M, \underline{R}, \mu)$ satisfying

$$\dot{R} > 0 \quad \text{in} \quad [T, T+d), \qquad R(T+d) - R(T) \geq \frac{1}{12}\underline{R}\mu - E_3.$$

where

$$\begin{aligned} E_3 := \Bigg| &\frac{\underline{R}}{12}\mu \frac{\frac{M^2}{4K} + \frac{M^2}{2\underline{R}K}}{\frac{1}{3}\underline{R}K\mu - \left(\frac{M^2}{6\underline{R}K} - \frac{M^2}{4K}\right)} \\ &+ \frac{1}{4K}\left(\frac{M^2}{6\underline{R}K} - \frac{M^2}{4K}\right)\left[1 - \frac{\frac{M^2}{4K} + \frac{M^2}{2\underline{R}K}}{\frac{1}{3}\underline{R}K\mu - \left(\frac{M^2}{6\underline{R}K} - \frac{M^2}{4K}\right)}\right] \\ &+ \left(\frac{M^2}{4K} + \frac{M^2}{2\underline{R}K}\right)\left[\frac{1}{4K}\log \frac{\frac{1}{3}\underline{R}K\mu - \left(\frac{M^2}{6\underline{R}K} - \frac{M^2}{4K}\right)}{\frac{M^2}{4K} + \frac{M^2}{2\underline{R}K}}\right]\Bigg|. \end{aligned}$$

Proof. Since the proof is rather lengthy, we leave it to Appendix C. □

6.2 A framework for the asymptotic lower bound of R

In this subsection, we study sufficient framework $(\mathcal{H})$ for the lower bounds of order parameter, and then present a rough estimate for the lower bound for R in Proposition 6.2:

$$\liminf_{t\to\infty} R(t) > \frac{2}{3}.$$

and then in the proof of Theorem 3.3 presented in next subsection, we will improve the above uniform lower bound by showing

$$\liminf_{t\to\infty} R(t) \geq 1 - \frac{|\mathcal{O}(1)|}{\sqrt{K}}.$$

We first list our main framework $(\mathcal{H})$ for the lower bound estimate of R as follows.

- $(\mathcal{H}1)$: The $\mathcal{C}^1$ initial data f_0 satisfies $R_0 > 0$, and g is compactly supported on the interval $[-M, M]$.

- $(\mathcal{H}2)$: The constant μ is sufficiently small and K is sufficiently large so that

$$\begin{aligned}
\frac{1}{2} &> 2\frac{M}{KR_0} + 4\frac{M}{KR_0^2} + \frac{2\sqrt{2}}{R_0\sqrt{R_0}}\sqrt{\frac{M}{K} + \mu},\\
K^2\mu &- \Big(\frac{2M^2}{R_0^2} - \frac{3M^2}{2R_0}\Big) > 0,\\
\text{and} \quad K &> \max\Big\{\frac{64M}{\sqrt{3}}, \frac{64\sqrt{3}M}{3-(\sqrt{3}-2R_0)^2}\Big\}.
\end{aligned}$$

- ($\mathcal{H}3$): The constant γ is contained in $(\frac{\pi}{3}, \frac{\pi}{2})$, and $E_i = E_i(K, M, \frac{R_0}{2}, \gamma),\ i = 1, 2$ and $E_3 = E_3(K, M, \frac{R_0}{2}, \mu)$ satisfy

$$R_0 - 2E_1 - E_2 > \frac{R_0}{2}, \qquad \mu\frac{R_0}{24} > 2E_1 + E_2 + E_3,$$

  where the positive constants E_1, E_2 and E_3 can be made sufficiently small by taking K sufficiently large and γ sufficiently close to $\frac{\pi}{2}$.

- ($\mathcal{H}4$): For a positive constant $\kappa \in (\frac{2}{3}, \frac{\sqrt{3}}{2})$, the coupling strength strength K is sufficiently large such that

$$\kappa < \frac{\sqrt{3}}{4} + \frac{1}{4}\sqrt{3 - \frac{64\sqrt{3}M}{K}}, \qquad \varepsilon_\kappa := \frac{\kappa+1}{\kappa^2}\frac{M}{K} + \frac{(1-\kappa)}{\kappa} < 1.$$

Under the above assumptions, we will derive

$$\liminf_{t\to\infty} R(t) \geq R_\infty := 1 + \frac{M}{K} - \sqrt{\frac{M^2}{K^2} + 4\frac{M}{K}}. \tag{6.14}$$

Thus, letting $K \to \infty$, the above estimate and $R \leq 1$ yield

$$\lim_{K\to\infty} \liminf_{t\to\infty} R(t) = 1.$$

Hence we obtain a complete phase synchronization in this asymptotic limit. Thus, the estimate (6.14) indicates the kinetic analogue of the practical synchronization estimates.

Now, we are ready to state the first proposition of this subsection. In this proposition, we show that the mass in $L^+_{\frac{\pi}{3}}$ remains above a constant, depending only on R_0 and E_1 in the set of times t where R is in non-increasing mode.

Proposition 6.1. *Suppose that the assumptions* $(\mathcal{H}1)$ - $(\mathcal{H}3)$ *hold, and assume that there exist* $t_0 \geq 0$ *and* $\underline{R} > 0$ *such that*

$$R(t_0) \geq R_0, \quad \inf_{0 \leq t \leq t_0} R(t) \geq \underline{R}.$$

Then, for any $t \geq t_0$ *satisfying* $\dot{R}(t) \leq 0$, *we have*

$$\mathcal{M}(L^{+}_{\frac{\pi}{3}}(t)) \geq \frac{1}{2}(R(t_0) + 1) - E_1. \tag{6.15}$$

Proof. Since the proof is rather lengthy, we postpone it to Apppendix D. □

Remark 6.2. *The estimate* (6.15) *looks similar to that appearing in Lemma 6.3, however in* (6.15), *we are comparing* $\mathcal{M}(L^{+}_{\frac{\pi}{3}}(t))$ *at any non-increasing instant* t *after* t_0 *with the fixed constant* $\frac{1}{2}(R(t_0) + 1) - E_1$. *Thus the estimate in Lemma 6.3 corresponds to a special situation of* (6.15).

Below, we present three corollaries followed by Proposition 6.1.

Corollary 6.1. *Suppose that assumptions* $(\mathcal{H}1)$ - $(\mathcal{H}3)$ *hold and there exists* $t_0 \geq 0$ *such that*

$$R(t_0) \geq R_0.$$

Then, we have

$$R(t) \geq R(t_0) - 2E_1 - E_2, \quad t \geq t_0 \quad \text{and} \quad \limsup_{t \to \infty} \dot{R}(t) \leq K\mu.$$

Proof. (i) First, assume $t_0 = 0$. Then, our hypotheses guarantee that the assumptions of Proposition 6.1 are satisfied. In the course of the proof of Proposition 6.1 in Appendix D, we have already shown that

$$R(t) \geq R(t_0) - 2E_1 - E_2, \quad \text{in } \mathcal{T}(t_0).$$

where

$$\mathcal{T}(t_0) := \Big\{ t \in [t_0, \infty) \ : \ \mathcal{M}(L^{+}_{\frac{\pi}{3}}(t^*)) \\ \geq \frac{1}{2}(R(t_0) + 1) - E_1 \quad \forall t^* \in [t_0, t] \cap \mathcal{N}(t_0) \Big\},$$

and we also showed $\mathcal{T} = [t_0, \infty)$. Thus, we have the desired estimate for $t_0 = 0$, and it follows from $(\mathcal{H}1)$ and $(\mathcal{H}3)$ that

$$R(t) > R_0 - 2E_1 - E_2 > \frac{R_0}{2} > 0, \quad t \in [0, \infty). \tag{6.16}$$

For the case $t_0 > 0$, note that by the above inequality, the assumptions $(\mathcal{H}1)$, $(\mathcal{H}2)$ and $(\mathcal{H}3)$, one can apply for Proposition 6.1 for $t_0 > 0$. Thus, the desired result follows by the same argument.

(ii) By definition of the order parameter, R is uniformly bounded, and

$$0 \leq \liminf_{t \to \infty} R(t) \leq 1.$$

Suppose that we have

$$\limsup_{t \to \infty} \dot{R} > K\mu. \tag{6.17}$$

Let $\varepsilon' > 0$ be sufficiently small so that

$$\frac{R_0}{24}\mu - \varepsilon' - 2E_1 - E_2 - E_3 > 0. \tag{6.18}$$

Such ε' exists by $(\mathcal{H}3)$. By the result in (i) we have

$$\liminf_{t\to\infty} R(t) \geq R_0 - 2E_1 - E_2. \tag{6.19}$$

On the other hand, by definition of $\liminf_{t\to\infty} R(t)$, there exists $t_0 \geq 0$, such that

$$R(t) \geq \liminf_{t\to\infty} R(t) - \varepsilon', \quad t \in [t_0, \infty). \tag{6.20}$$

Thanks to (6.17), boundedness of R and continuity of $\dot{R}$ in Lemma 6.2, there exists $t_1, t_2 \geq t_0$ such that

$$\dot{R}(t_1) > K\mu \quad \text{and} \quad \dot{R}(t_2) = K\mu.$$

Thus, it follows from Lemma 6.5, (6.20), (6.19) and (6.18) that there exists $d > 0$ such that

$$\begin{aligned} R(t_2 + d) &\geq R(t_2) + \frac{R_0}{24}\mu - E_3 \\ &\geq \liminf_{t\to\infty} R(t) + \frac{R_0}{24}\mu - \varepsilon' - E_3 \\ &\geq R_0 + \frac{R_0}{24}\mu - \varepsilon' - 2E_1 - E_2 - E_3 \\ &\geq R_0. \end{aligned}$$

Then, by the third inequality as above and (6.16), we have

$$R(t) \geq \liminf_{t\to\infty} R(t) + \frac{R_0}{24}\mu - \varepsilon' - 2E_1 - E_2 - E_3, \quad \text{in } [t_2 + d, \infty).$$

We use (6.18) to see that the above expression contradicts the definition of $R_* = \liminf_{t\to\infty} R$. This yields the desired result. $\square$

We next show that the L^2 norm of f in an interval of length $\frac{\pi}{3}$, centered at $-\phi(t)$, decays exponentially after some time. This is analogous to the phenomenon in (4.15).

Corollary 6.2. *Suppose that assumptions* $(\mathcal{H}1)$ - $(\mathcal{H}3)$ *hold. Then, there exists* $T \geq 0$ *such that*

$$\begin{aligned} &\iint_{L^-_{\frac{\pi}{3}}(t)\times\mathbb{R}} |f|^2 \; d\theta d\omega \leq e^{-\frac{KR(0)}{4}(t-T)} \iint_{L^-_{\frac{\pi}{3}}(T)\times\mathbb{R}} |f(T)|^2 \; d\theta d\omega, \quad t \geq T, \\ &\mathcal{M}\big(L^-_{\frac{\pi}{3}}(t)\big) \leq \left(\frac{\pi}{3}\right)^{\frac{1}{2}} e^{-\frac{KR(0)}{8}(t-T)} \sqrt{\iint_{L^-_{\frac{\pi}{3}}(T)\times\mathbb{R}} f^2(\theta,\omega,T) \; d\theta d\omega.} \end{aligned} \tag{6.21}$$

Proof. We postpone its proof in Appendix E. □

Before we present the last proposition showing that R will remain above $\frac{2}{3}$ after some time, we present a preparatory result.

Lemma 6.6. *Suppose that assumptions* $(\mathcal{H}1)$ - $(\mathcal{H}3)$ *hold. Then, the following estimate holds.*

$$\begin{aligned} \frac{2R(t)-\sqrt{3}+2\sqrt{3}\mathcal{M}(L^-_{\frac{\pi}{3}}(t))}{2-\sqrt{3}} &\leq \mathcal{M}(L^+_{\frac{\pi}{3}}(t)) \\ &\leq \frac{2R(t)+\sqrt{3}+\mathcal{M}(L^-_{\frac{\pi}{3}}(t))(2-\sqrt{3})}{2\sqrt{3}}, \quad t \in [0,\infty). \end{aligned}$$

Proof. Note that

$$\begin{aligned}
R(t) &= \iint_{\mathbb{T}\times\mathbb{R}} \langle e^{\mathrm{i}\theta}, e^{\mathrm{i}\phi}\rangle f \ d\theta d\omega \\
&= \iint_{\left(\mathbb{T}\setminus(L^+_{\frac{\pi}{3}}(t)\cup L^-_{\frac{\pi}{3}}(t))\right)\times\mathbb{R}} \langle e^{\mathrm{i}\theta}, e^{\mathrm{i}\phi}\rangle f \ d\theta d\omega + \iint_{L^+_{\frac{\pi}{3}}(t)\times\mathbb{R}} \langle e^{\mathrm{i}\theta}, e^{\mathrm{i}\phi}\rangle f \ d\theta d\omega \\
&+ \iint_{L^-_{\frac{\pi}{3}}(t)\times\mathbb{R}} \langle e^{\mathrm{i}\theta}, e^{\mathrm{i}\phi}\rangle f \ d\theta d\omega.
\end{aligned} \tag{6.22}$$

• (Lower bound estimate): We use similar arguments in Lemma 6.2 to see that

$$\begin{aligned}
R(t) &\ge -\sin(\frac{\pi}{3})(1-\mathcal{M}(L^+_{\frac{\pi}{3}}(t))-\mathcal{M}(L^-_{\frac{\pi}{3}}(t))+\sin\frac{\pi}{3}\mathcal{M}(L^+_{\frac{\pi}{3}}(t))-\mathcal{M}(L^-_{\frac{\pi}{3}}(t)) \\
&= \sqrt{3}\mathcal{M}(L^+_{\frac{\pi}{3}}(t))-\frac{\sqrt{3}}{2}-\mathcal{M}(L^-_{\frac{\pi}{3}}(t))\Big(1-\frac{\sqrt{3}}{2}\Big).
\end{aligned}$$

• (Upper bound estimate): We use again (6.22) to obtain

$$\begin{aligned}
R(t) &\le \sin\frac{\pi}{3}(1-\mathcal{M}(L^+_{\frac{\pi}{3}}(t))-\mathcal{M}(L^-_{\frac{\pi}{3}}(t))+\mathcal{M}(L^+_{\frac{\pi}{3}}(t))-\sin\frac{\pi}{3}\mathcal{M}(L^-_{\frac{\pi}{3}}(t)) \\
&= \frac{\sqrt{3}}{2}+\mathcal{M}(L^+_{\frac{\pi}{3}}(t))\Big(1-\frac{\sqrt{3}}{2}\Big)-\sqrt{3}\mathcal{M}(L^-_{\frac{\pi}{3}}(t)).
\end{aligned}$$

This yields the desired inequalities. □

Before we state our proposition, we introduce several quantities: for every $T \ge 0$ and $\eta > 0$, we define $\beta_{T,\eta} \ : [T,\infty) \to \mathbb{R}$ as a solution of the inhomogeneous Riccati ODE:

$$\begin{aligned}
&\dot{\beta}_{T,\eta} = \frac{K}{4\sqrt{3}}\Big(-\beta^2_{T,\eta}+\frac{\sqrt{3}}{2}\beta_{T,\eta}-\frac{4\sqrt{3}M}{K}-\eta\Big), \quad t > T, \\
&\beta_{T,\eta}(T) = R(T),
\end{aligned} \tag{6.23}$$

and we set the solutions of the following quadratic equation as $r_{\pm}(\eta)$:

$$x^2 - \frac{\sqrt{3}}{2}\, x + 2\frac{4\sqrt{3}M}{K} + \eta = 0.$$

More precisely, we have

$$r_-(\eta) := \frac{\sqrt{3}}{4} - \frac{1}{2}\sqrt{\frac{3}{4} - \frac{16\sqrt{3}M}{K} - 4\eta},$$

$$r_+(\eta) := \frac{\sqrt{3}}{4} + \frac{1}{2}\sqrt{\frac{3}{4} - \frac{16\sqrt{3}M}{K} - 4\eta}.$$

Note that if $\beta_{T,\eta}(T) > r_-(\eta)$, then we have

$$\lim_{t\to\infty} \beta_{T,\eta}(t) = r_+(\eta). \tag{6.24}$$

Proposition 6.2. *Suppose that the assumptions* $(\mathcal{H}1)$ - $(\mathcal{H}4)$ *hold. Then, we have*

$$\liminf_{t\to\infty} R(t) \geq r_+(0) > \kappa > \frac{2}{3}. \tag{6.25}$$

Proof. Let ε_0 be a small positive constant satisfying

$$\frac{R_0}{2} > r_-(\varepsilon_0). \tag{6.26}$$

Note that the assumption $(\mathcal{H}2)$ on K implies

$$K > \max \frac{64\sqrt{3}M}{3 - (\sqrt{3} - 2R_0)^2} \quad \Rightarrow \quad \frac{R_0}{2} > r_-(0).$$

Thus, the continuity of $r_-(\cdot)$ with respect to its argument implies the existence of ε_0. We set η and d be positive numbers satisfying

$$\eta < \varepsilon_0, \qquad d > \frac{8}{KR_0} \log\left[\frac{1}{\eta}\Big(1 + \frac{\sqrt{3}}{2}\Big)\Big(\frac{\pi}{3}\Big)^{\frac{1}{2}} \sqrt{\iint_{L^-_{\frac{\pi}{3}}(T)\times\mathbb{R}} f^2(\theta,\omega,T)\; d\theta d\omega}\,\right]. \tag{6.27}$$

Here T is the one given by Corollary 6.2.

By Lemma 5.2, we have

$$\dot{R} = -\iint_{\mathbb{T}\times\mathbb{R}} \sin(\theta-\phi)\omega f(\theta,\omega,t)\ d\theta d\omega$$

$$+KR\iint_{\mathbb{T}\times\mathbb{R}} \sin^2(\theta-\phi) f(\theta,\omega,t)\ d\theta d\omega$$

$$\geq -M + \frac{KR}{4}\Big(1-\mathcal{M}(L^{+}_{\frac{\pi}{3}})-\mathcal{M}(L^{-}_{\frac{\pi}{3}})\Big).$$

Then, we use Lemma 6.6 to get

$$\begin{aligned}
\dot{R} &\geq -M + \frac{KR}{4}\Big[1-\frac{2R+\sqrt{3}+\mathcal{M}(L^{-}_{\frac{\pi}{3}})(2-\sqrt{3})}{2\sqrt{3}} - \mathcal{M}(L^{-}_{\frac{\pi}{3}})\Big]\\
&= -M + \frac{KR}{8} - \frac{KR^2}{4\sqrt{3}} - \frac{KR}{4}\Big(\frac{2+\sqrt{3}}{2\sqrt{3}}\Big)\mathcal{M}(L^{-}_{\frac{\pi}{3}})\\
&= \frac{K}{4\sqrt{3}}\Big[-R^2+\frac{\sqrt{3}}{2}R-\frac{4\sqrt{3}M}{K} - R\Big(1+\frac{\sqrt{3}}{2}\Big)\mathcal{M}(L^{-}_{\frac{\pi}{3}})\Big]\\
&\geq \frac{K}{4\sqrt{3}}\Big[-R^2+\frac{\sqrt{3}}{2}R-\frac{4\sqrt{3}M}{K} - \Big(1+\frac{\sqrt{3}}{2}\Big)\mathcal{M}(L^{-}_{\frac{\pi}{3}})\Big].
\end{aligned}$$

Hence, it follows from (6.21) and (6.27) that we have

$$\dot{R} \geq \frac{K}{4\sqrt{3}}\Big(-R^2+\frac{\sqrt{3}}{2}R-\frac{4\sqrt{3}M}{K}-\eta\Big) \quad \text{in} \quad [T+d,\infty). \tag{6.28}$$

Moreover, thanks to Corollary 6.1, we also see that

$$R(t) > \frac{R_0}{2}, \qquad \forall\ t>0.$$

We use (6.26) to obtain

$$R(T+d) > r_-(\eta).$$

Thus, (6.23) and (6.28) yield

$$R \geq \beta_{T+d,\eta} \quad \text{in} \quad [T+d, \infty).$$

Since

$$R(T+d) \geq \beta_{T+d,\eta}(T+d) > r_-(\eta),$$

we get the desired result from (6.24) and the fact that η was arbitrary small.

$\square$

Remark 6.3. *The lower bound for* $\liminf_{t\to\infty} R(t)$ *in* (6.25) *will be improved in Theorem 3.3 so that* $\liminf\limits_{t\to\infty} R(t) \to 1$, *as* $K \to \infty$.

6.3 Proof of Theorem 3.3

In this subsection, we will provide the proof for Theorem 3.3. For this purpose, for κ, ε and δ in $(0,1)$ and $t \geq 0$, we define the sets:

$$\begin{aligned} Y_{\kappa,\varepsilon}(t) &:= \left\{ \theta \in \mathbb{T} \;:\; \cos(\theta - \phi(t)) > \sqrt{1-\varepsilon_\kappa^2} - \varepsilon \right\}, \\ \tilde{Y}_\delta(t) &:= \left\{ \theta \in \mathbb{T} \;:\; \cos(\theta - \phi(t)) \leq -\delta \right\}, \end{aligned}$$

where for $\kappa \in (0,1)$, ε_κ and T_κ are positive constants satisfying the following relations, respectively:

$$\varepsilon_\kappa := \frac{\kappa+1}{\kappa^2}\frac{M}{K} + \frac{(1-\kappa)}{\kappa} < 1 \quad \text{and} \quad R > \kappa \quad \text{in} \quad [T_\kappa, \infty).$$

Lemma 6.7. *Suppose that for some positive constants κ and δ in $(0,1)$, the order parameter R satisfies*

$$R > \underline{R} \quad \text{in} \quad [0,\infty), \qquad R > \kappa \quad \text{in} \quad [T_\kappa,\infty), \qquad \delta < \sqrt{1-\varepsilon_\kappa^2},$$

and we also assume that f_0 is bounded and supp $f_0(\cdot,\omega) \subset [-M,M]$. *Then we have*

$$\lim_{t\to 0} ||f(t)\mathbb{1}_{\tilde{Y}_\delta(t)}||_{L^\infty(\mathbb{T}\times[-M,M])} = 0,$$

where the positive constant ε_κ is defined in ($\mathcal{H}4$).

Proof. By Lemma 6.1, we have

$$||f(T_\kappa)||_{L^\infty(\mathbb{T}\times[-M,M])} \leq ||f_0||_{L^\infty(\mathbb{T}\times[-M,M])} e^{KT_\kappa}. \tag{6.29}$$

Choose constants θ^*, ω^*, and t^* satisfying

$$\omega^* \in [-M,M] \quad \text{and} \quad \cos(\theta^* - \phi(t^*)) \leq -\delta.$$

By Corollary F.1 in the appendix, we know that the characteristic $(\theta(t),\omega^*)$ through θ^* and ω^* at t^* satisfies

$$\cos(\theta(t) - \phi(t)) \leq -\delta, \quad \forall\ t \in [T_\kappa, t^*).$$

Thus, proceeding as in (6.1), we get

$$\begin{aligned}\frac{d}{dt} f(\theta(t),\omega^*,t) &= KR(t)\cos(\theta(t)-\phi(t)) f(\theta(t),\omega^*,t)\\ &\leq -K\kappa\delta f(\theta(t),\omega^*,t), \qquad t \in [T_\kappa, t^*).\end{aligned}$$

We use (6.29) and Gronwall's inequality to obtain

$$f(\theta^*, \omega^*, t^*) = f(\theta(t^*), \omega(t^*), t^*) \leq ||f_0||_{L^\infty(\mathbb{T}\times[-M,M])} e^{KT_\kappa} e^{-K\kappa\delta(t^*-T_\kappa)}$$

Since (θ^*, ω^*) was an arbitrary point in $\tilde{Y}_\delta(t^*) \times [-M, M]$, we have

$$||f(t^*)\mathbb{1}_{\tilde{Y}_\delta(t^*)}||_{L^\infty(\mathbb{T}\times[-M,M])} \leq ||f_0||_{L^\infty(\mathbb{T}\times[-M,M])} e^{-K\kappa\delta t^*} e^{K(1+\kappa\delta)T_\kappa}.$$

By letting $t^* \to \infty$, we obtain the desired result. □

Proposition 6.3. *Suppose that the following conditions hold.*

1. *The order parameter satisfies*

$$R > \underline{R} \quad \text{in} \quad [0, \infty), \quad R > \kappa \quad \text{in} \quad [T_\kappa, \infty) \quad \textit{and} \quad \varepsilon_\kappa < 1,$$

for some positive constants κ and ε_κ.

2. *f_0 is bounded and supp $f_0(\cdot, \omega) \subset [-M, M]$.*

Then, we have

$$\lim_{t\to 0} ||f(t)\mathbb{1}_{\mathbb{T}\setminus Y_{\kappa,\varepsilon}(t)}||_{L^\infty(\mathbb{T}\times[-M,M])} = 0, \quad \forall \ \varepsilon \in (0, \sqrt{1-\varepsilon_\kappa^2}).$$

Proof. Let ε and δ be positive constants such that

$$\varepsilon < \sqrt{1-\varepsilon_\kappa^2} \qquad \text{and} \qquad \delta < \sqrt{1-\varepsilon_\kappa^2} - \varepsilon.$$

By Lemma 6.7, for any $\varepsilon' > 0$, there exists $T_{\varepsilon'}$ such that

$$R(t) > \kappa \quad \text{and} \quad \|f_t \mathbb{1}_{\tilde{Y}_\delta(t)}\|_{L^\infty(\mathbb{T}\times[-M,M])} < \varepsilon', \quad t \in [T_{\varepsilon'}, \infty). \tag{6.30}$$

For given ε and κ, we define

$$F_\kappa(q) := \kappa K\Big(\sqrt{1-q^2} - \varepsilon_\kappa\Big)\sqrt{1-q^2}, \quad q \in [-1,1],$$

$$D(\varepsilon,\kappa) := \frac{2\big(\sqrt{1-\varepsilon_\kappa^2} - \varepsilon\big)}{F_\kappa\Big(\sqrt{1-\varepsilon_\kappa^2} - \varepsilon\Big)}.$$

Let θ^*, ω^*, and t^* have the property that

$$\omega^* \in [-M, M], \qquad t^* \geq T_{\varepsilon'} + D(\varepsilon,\kappa), \quad \cos(\theta^* - \phi(t^*)) \leq \sqrt{1-\varepsilon_\kappa^2} - \varepsilon.$$

Let $(\theta(t), \omega^*, t)$ be the forward characteristic through $(\theta^*, \omega^*, t^*)$ at time t^*. By Corollary F.1, we have

$$\cos(\theta(t^*-d) - \phi(t^*-d)) \leq -\sqrt{1-\varepsilon_\kappa^2} + \varepsilon \leq -\delta, \quad \text{where} \quad d < D(\varepsilon,\kappa).$$

Proceeding as in (6.1), we get

$$\begin{aligned}\frac{d}{dt} f(\theta(t), \omega^*, t) &= KR(t)\cos(\theta(t) - \phi(t)) f(\theta(t), \omega^*, t) \\ &\leq K f(\theta(t), \omega^*, t), \quad t \in [T_\kappa, t^*).\end{aligned}$$

By Gronwall's lemma, we obtain

$$f(\theta^*, \omega^*, t^*) = f(\theta(t^*), \omega(t^*), t^*) \leq e^{Kd} f(\theta(t^*-d), \omega(t^*-d), t^*) \leq e^{KD(\varepsilon,\kappa)}\varepsilon',$$

where in the last line we have used (6.30) and the fact that $\theta(t^*-d)$ contained in $\tilde{Y}_\delta(t^*-d)$.

Since (θ^*, ω^*) was an arbitraty point in $\mathbb{T}\backslash Y_{\kappa,\varepsilon}(t^*)$, where

$$Y_{\kappa,\varepsilon}(t^*) := \big\{\theta \in \mathbb{T} \ : \ \cos(\theta - \phi(t^*)) > \sqrt{1-\varepsilon_\kappa^2} - \varepsilon'\big\},$$

we conclude

$$||f(t^*)\mathbb{1}_{\mathbb{T}\setminus Y_{\kappa,\varepsilon}(t^*)}||_{L^\infty(\mathbb{T}\times[-M,M])} \leq e^{KD(\varepsilon,\kappa)}\varepsilon'.$$

Additionally, since t^* was an arbitrary point in $[T_{\varepsilon'} + D(\varepsilon,\kappa),\infty)$, we get

$$||f(t)\mathbb{1}_{\mathbb{T}\setminus Y_{\kappa,\varepsilon}(t)}||_{L^\infty(\mathbb{T}\times[-M,M])} \leq e^{KD(\varepsilon,\kappa)}\varepsilon', \quad \forall t \in [T_{\varepsilon'} + D(\varepsilon,\kappa),\infty).$$

Since ε' is an arbitrary positive number, we have the desired estimate:

$$\lim_{t\to\infty} ||f(t)\mathbb{1}_{\mathbb{T}\setminus Y_{\kappa,\varepsilon}(t)}||_{L^\infty(\mathbb{T}\times[-M,M])} = 0.$$

$\square$

We are now ready to provide the proof of Theorem 3.3. Suppose that the assumptions $(\mathcal{H}1)$ to $(\mathcal{H}4)$ hold, and we assume

$$\frac{4}{15}\left(\sqrt{1-\frac{\sqrt{2}}{2}}-\frac{1}{2}\right) > \frac{M}{K} \quad \text{or equivalently}$$

$$K > \frac{15M}{2(\sqrt{4-2\sqrt{2}}-1)}. \tag{6.31}$$

Then, Theorem 3.3 follows once we prove the following claims:

1. $$\liminf_{t\to\infty} R(t) \geq R_\infty := 1 + \frac{M}{K} - \sqrt{\frac{M^2}{K^2} + 4\frac{M}{K}}.$$

2. There exists a time dependent interval L_∞, centered at $\phi(t)$, with constant width

$$2\ \arccos\left(\sqrt{1-\left[\frac{M}{K}\frac{(1+R_\infty)}{R_\infty^2}+\frac{1-R_\infty}{R_\infty}\right]^2}\right)+\varepsilon,$$

such that

$$\lim_{t\to\infty} ||f(t)\mathbb{1}_{\mathbb{T}\setminus L_\infty(t)}||_{L^\infty(\mathbb{T}\times[-M,M])} = 0, \quad \text{for every } \varepsilon \text{ in } (0,1).$$

As $K \to \infty$, R_∞ tends to 1 and the width of L_∞ can be made arbitrarily small.

Proof of claim: Note that the second item (2) follows from the item (1) by apply Proposition 6.3. To prove the item (1), by the assumptions $(\mathcal{H}1)$ - $(\mathcal{H}3)$ and Corollary 6.1, we have

$$||f_0||_{L^\infty(\mathbb{T}\times[-M,M])} < \infty \quad \text{and} \quad R(t) > \frac{R_0}{2} \quad \forall t \in [0,\infty).$$

Second, Proposition 6.2 and $(\mathcal{H}4)$ yield

$$R > \frac{2}{3}, \quad \text{in } [T,\infty) \ \text{ for some } T \geq 0.$$

Now, we improve this lower bound. Let $\bar{\varepsilon}_0 > 0$ be sufficiently small so that

$$\frac{4}{15}\left(\sqrt{1-\frac{\sqrt{2}}{2}-\bar{\varepsilon}_0}-\frac{1}{2}\right) > \frac{M}{K}. \tag{6.32}$$

Such $\bar{\varepsilon}_0$ exists by assumption (6.31). Moreover, let η be an arbitrary positive constant in $(0,1)$. By Proposition 6.3 and the assumption $(\mathcal{H}4)$,

$$\lim_{t\to 0} ||f(t)\mathbb{1}_{\mathbb{T}\setminus Y_{\frac{2}{3},\eta}(t)}||_{L^\infty(\mathbb{T}\times[-M,M])} = 0.$$

From this, we also see

$$\lim_{t\to\infty} \mathcal{M}\left(\left\{\theta \in \mathbb{T} \ : \ \cos(\theta-\phi(t)) \leq \sqrt{1-\left(\frac{2/3+1}{(2/3)^2}\frac{M}{K}+\frac{1-2/3}{2/3}\right)^2}-\eta\right\}\right) = 0, \tag{6.33}$$

or equivalently,

$$\lim_{t\to\infty} \mathcal{M}\left(\left\{\theta \in \mathbb{T}: \ \cos(\theta - \phi(t)) \le \sqrt{1 - \left(\frac{15}{4}\frac{M}{K} + \frac{1}{2}\right)^2} - \eta\right\}\right) = 0.$$

Then, using (6.32) in the previous limit, we deduce

$$\lim_{t\to\infty} \mathcal{M}(A(t)) = 0,$$

where

$$A(t) := \left\{\theta \in \mathbb{T}: \ \cos(\theta - \phi(t)) \le \sqrt{\frac{\sqrt{2}}{2} + \bar{\varepsilon}_0} - \eta\right\}. \tag{6.34}$$

Since

$$\begin{aligned} R(t) &= \iint_{(\mathbb{T}\setminus A(t))\times\mathbb{R}} \langle e^{\mathrm{i}\phi(t)}, e^{\mathrm{i}\theta}\rangle f \ d\theta d\omega + \iint_{A(t)\times\mathbb{R}} \langle e^{\mathrm{i}\phi(t)}, e^{\mathrm{i}\theta}\rangle f \ d\theta d\omega \\ &\ge \Big(1 - \mathcal{M}(A(t))\Big)\left(\sqrt{\frac{\sqrt{2}}{2} + \bar{\varepsilon}_0} - \eta\right) - \mathcal{M}(A(t)), \end{aligned} \tag{6.35}$$

and η can be made arbitrarely small, we obtain

$$R_* := \liminf_{t\to\infty} R(t) \ge \sqrt{\frac{\sqrt{2}}{2} + \bar{\varepsilon}_0}. \tag{6.36}$$

Then, there exists T_0 such that

$$R(t) \ge \sqrt{\frac{\sqrt{2}}{2}} \quad \forall t \in [T_0, \infty).$$

We next claim:

$$R_* \ge 1 + \frac{M}{K} - \sqrt{\frac{M^2}{K^2} + 4\frac{M}{K}} =: \alpha_-(M, K).$$

Suppose not, i.e.,

$$R_* < 1 + \frac{M}{K} - \sqrt{\frac{M^2}{K^2} + 4\frac{M}{K}}.$$

Then, by (6.36) we have

$$R_* \in \left(\sqrt{\frac{\sqrt{2}}{2}},\ 1+\frac{M}{K}-\sqrt{\frac{M^2}{K^2}+4\frac{M}{K}}\right). \tag{6.37}$$

Since the roots of the polynomial

$$Q(x) = x^2 - 2\left(1+\frac{M}{K}\right)x + 1 - 2\frac{M}{K},$$

are

$$\alpha_-(K,M) = 1+\frac{M}{K}-\sqrt{\frac{M^2}{K^2}+4\frac{M}{K}}, \qquad \alpha_+(K,M) := 1+\frac{M}{K}+\sqrt{\frac{M^2}{K^2}+4\frac{M}{K}},$$

we have

$$Q(x) > 0 \quad \text{in} \quad (-\infty, \alpha_-(K,M)).$$

It follows from (6.37) that

$$0 < R_*^2 - 2(1+\frac{M}{K})R_* + 1 - 2\frac{M}{K}, \quad \text{or equivalently} \quad 0 < \frac{1}{2}R_*^2 - \left(1+\frac{M}{K}\right)R_* + \frac{1}{2} - \frac{M}{K}.$$

This yields

$$\frac{1}{2}(R_*^2 - 1) < -R_*(1-R_*) - \frac{M}{K}(1+R_*). \tag{6.38}$$

By (6.31) and (6.37) we have

$$0 < R_*(1-R_*) + \frac{M}{K}(1+R_*) < \frac{1}{4} + 2\frac{M}{K} < 1.$$

We again use (6.38) to see

$$\begin{aligned}\left(\sqrt{\frac{\sqrt{2}}{2}}\right)^4 \left(R_*^2 - 1\right) = \frac{1}{2}(R_*^2 - 1) &< -R_*(1-R_*) - \frac{M}{K}(1+R_*) \\ &< -\left[R_*(1-R_*) + \frac{M}{K}(1+R_*)\right]^2.\end{aligned}$$

By the above expression and (6.37),

$$R_*^4\big(R_*^2-1\big) < \left(\sqrt{\frac{\sqrt{2}}{2}}\right)^4\big(R_*^2-1\big) < -\left[R_*(1-R_*)+\frac{M}{K}(1+R_*)\right]^2.$$

Hence, we have

$$R_* < \sqrt{1-\left[\frac{M}{K}\frac{(1+R_*)}{R_*^2}+\frac{1-R_*}{R_*}\right]^2}. \tag{6.39}$$

We set

$$R_{*,\varepsilon'} := R_* - \varepsilon'.$$

By construction, there exists $T_{\varepsilon'}$ such that

$$R(t) \geq R_{*,\varepsilon'} \quad \text{in} \quad [T_{\varepsilon'}, \infty).$$

Moreover, it follows from Proposition 6.3 that we have

$$\lim_{t\to\infty} \mathcal{M}\big(\mathbb{T}\backslash Y_{R_{*,\varepsilon'},\varepsilon}(t)\big) = 0, \quad \text{for any } \varepsilon > 0.$$

Then, by the same arguments as in (6.35), we get

$$\begin{aligned} R(t) \geq \big(1-\mathcal{M}(\mathbb{T}\backslash Y_{R_{*,\varepsilon'},\varepsilon}(t)\big)\Bigg[&\sqrt{1-\left[\frac{M}{K}\frac{(1+R_{*,\varepsilon'})}{R_{*,\varepsilon'}^2}+\frac{1-R_{*,\varepsilon'}}{R_{*,\varepsilon'}}\right]^2} \\ &- \varepsilon\Bigg] - \mathcal{M}(\mathbb{T}\backslash Y_{R_{*,\varepsilon'},\varepsilon}(t)). \end{aligned} \tag{6.40}$$

Thus, we have

$$R_* \geq \sqrt{1-\left[\frac{M}{K}\frac{(1+R_{*,\varepsilon'})}{R_{*,\varepsilon'}^2}+\frac{1-R_{*,\varepsilon'}}{R_{*,\varepsilon'}}\right]^2} - \varepsilon.$$

Since ε and ε' can be made arbitrarily small, using inequality (6.39) and the above expression we get

$$R_* \geq \sqrt{1 - \left[\frac{M}{K}\frac{(1+R_*)}{R_*^2} + \frac{1-R_*}{R_*}\right]^2} > R_*,$$

which yields a contradiction. Therefore, we have

$$R_* \geq 1 + \frac{M}{K} - \sqrt{\frac{M^2}{R_*^2} + 4\frac{K}{M}}.$$

This completes the proof of Theorem 3.3.

Remark 6.4. *We now briefly discuss how the lower bound on $\kappa > \frac{2}{3}$ in* ($\mathcal{H}3$) *was determined. In the above proof, we used the property*

$$A(t) \subset \left\{\theta \in \mathbb{T} : \cos(\theta - \phi(t)) \leq \sqrt{1 - \Big(\frac{\kappa+1}{\kappa^2}\frac{M}{K} + \frac{1-\kappa}{\kappa}\Big)^2} - \eta\right\}$$

in (6.33) *and* (6.34)*, which is equivalent to*

$$1 - \Big(\frac{\kappa+1}{\kappa^2}\frac{M}{K} + \frac{1-\kappa}{\kappa}\Big)^2 > \frac{\sqrt{2}}{2} + \bar{\varepsilon}_0. \tag{6.41}$$

Thus, if the following inequality holds,

$$1 - \Big(\frac{1-\kappa}{\kappa}\Big)^2 > \frac{\sqrt{2}}{2}, \tag{6.42}$$

(6.41) *is satisfied for small $\bar{\varepsilon}_0$ and sufficiently large K (depending on $\bar{\varepsilon}_0$). Since the inequality* (6.42) *is equivalent to*

$$\sqrt{2} - \sqrt{2-\sqrt{2}} < \kappa < \sqrt{2} + \sqrt{2-\sqrt{2}},$$

and $\sqrt{2} - \sqrt{2-\sqrt{2}} \approx 0.6488$, we choose $\kappa > \frac{2}{3}$ for the simplicity.

7 Conclusion

In this paper, we have presented several results on the asymptotic dynamics of the Kuramoto-Sakaguchi equation which is obtained from the Kuramoto model in the mean-field limit. For a large ensemble of Kuramoto oscillators, it is very expensive to study the dynamics of the oscillators directly via the Kuramoto model. So, from the beginning of the study on Kuramoto oscillators, the corresponding mean-field model, namely the Kuramoto-Sakaguchi equation has been widely used in the physics literature for the phase transition phenomena of large ensembles of Kuramoto oscillators. For example, Kuramoto himself employed a self-consistent theory based on the linearized Kuramoto-Sakaguchi equation, to derive a critical coupling strength for the phase transition from disordered states to partially ordered states (see [1]). However, existence of steady states and chimera states, as well as their non-linear stability are still far from complete understanding. In this long paper, we have studied phase concentration in a large coupling regime, for a large ensemble of oscillators. First, in the identical natural frequency case, we showed that mass of the ensemble concentrates exponentially fast at the average phase. In particular, the mass on each interval containing the average phase is non-decreasing over time, whereas the mass outside the interval decays to zero asymptotically. This illustrates the formation of a point cluster for the large ensemble of Kuramoto oscillators, which is a stable solution. It is interesting to note that, on the other hand, the Kuramoto model allows the unstable bi-polar state as an asymptotic pattern. Second, for the non-identical natural frequencies, i.e., the general case, we showed that the phases of a large ensemble of Kuramoto oscillators will aggregate inside a small interval around the average phase as the coupling strength increases. This is a similar feature as in the finite-dimensional Kuramoto model. Our third result is a quantitative lower bound for the amplitude order parameter. From a series of technical lemmata, we obtain an asymptotic formula for the amplitude order parameter in a large coupling strength regime, which also shows that a point cluster pattern arises as the coupling strength becomes sufficiently large. There are still lots of issues to be resolved on the large-time dynamics of the Kuramoto-Sakaguchi equation. To name a few, we mention three outstanding problems. First, we have not yet shown the existence of stationary solutions for the Kuramoto-Sakaguchi equation. Thus, our present results can be a first foot step toward this direction. Second, our estimates on the ensemble of Kuramoto oscillators with distributed natural frequencies are strongly relying on the large coupling strength. In particular, we have not optimized the size of coupling strength. Thus, one interesting question is to find the critical coupling strength for the phase transition from the partially ordered state to the fully ordered state (complete synchronization). Third, it will be also interesting to investigate the intermediate regime where the coupling strength is not too small nor too large, especially, regarding existence and stability of partially synchronized states. These issues will be addressed in future work.

Acknowledgement

The work of S.-Y. Ha is partially supported by a National Research Foundation of Korea Grant (2014R1A2A2A05002096) funded by the Korea government, and the work of J. Park was supported by NRF(National Research Foundation of Korea) Grant funded by the Korean Government(NRF-2014-Fostering Core Leaders of the Future Basic Science Program) The work of Y.-H.Kim is partially supported by Natural Sciences and Engineering Research Council of Canada Discovery Grants 371642-09 and 2014-05448 as well as the Alfred P. Sloan Research Fellowship 2012–2016. Part of this research has been done while the authors were participating in the fall Semester 2015 in Analysis at École Normale Supérior de Lyon (ENS-Lyon), France. We are grateful for the hospitality of ENS-Lyon and Prof. Albert Fathi.

Appendix

Appendix

A Otto calculus

In this section, we review the Otto calculus dealing with gradient flows on the Wasserstein space, and explain how the K-S equation can be regarded as a gradient flow on the Wasserstein space.

A.1 The K-S equation as a gradient flow

We now formulate the gradient flow for the potential V_k via the Otto calculus, and see that it coincides with the K-S equation (4.1), equivalently (2.21) for the identical oscillators case. This is a rather standard procedure (see [93], for example). For this, we first recall the Otto calculus introduced in [83], which gives a formal Riemannian metric on the space of absolutely continuous probability measures $\mathbb{P}_{ac}(\mathbb{M})$ on a Riemannian manifold $\mathbb{M}$.

Consider two curves $\rho^1, \rho^2 : (-\varepsilon, \varepsilon) \to \mathbb{P}_{ac}(\mathbb{M})$ with the common value ρ at $t = 0$, i.e. $\rho^1(0) = \rho^2(0) = \rho$. Assume that they are differentiable and the Riemmanian product between the time derivatives $\dot{\rho}^1(0), \dot{\rho}^2(0)$ is given by

$$\langle \dot{\rho}^1(0), \dot{\rho}^2(0) \rangle_{W_2} = \int_{\mathbb{M}} \langle \nabla\varphi^1\big|_{t=0}, \nabla\varphi^2\big|_{t=0} \rangle \rho \quad d\mathrm{vol},$$

where the bracket $\langle \cdot, \cdot \rangle$ is the Riemannian product on the underlying space of $\mathbb{M}$, and the functions φ^1 and φ^2 are determined by solving the continuity equation:

$$\frac{\partial}{\partial t}\rho^i + \mathrm{div}(\rho^i \nabla\varphi^i) = 0, \quad i = 1, 2.$$

With respect to this metric $\langle \cdot, \cdot \rangle_{W_2}$, the gradient of a given functional F : $\mathbb{P}_{ac}(\mathbb{M}) \to \mathbb{R}$, can be considered as a vector field, denoted by $\mathrm{grad}_\rho F$ on $\mathbb{M}$, such that for a one-parameter differentiable family $\tau \mapsto g_\tau \in \mathbb{P}_{ac}(\mathbb{M})$ with $g_0 = \rho$, satisfies

$$\frac{d}{dt}\bigg|_{\tau=0} F(g_\tau) = \int_{\mathbb{M}} \langle \nabla\varphi\big|_{\tau=0}, \mathrm{grad}_\rho F \rangle \rho \ d\mathrm{vol},$$

where the vector field $\nabla\varphi$ solves the continuity equation:

$$\frac{\partial g_t}{\partial t} + \mathrm{div}(g_t \nabla\varphi) = 0. \tag{A.1}$$

Then, the gradient flow of F is a one-parameter family $t \mapsto \rho_t \in \mathbb{P}_{ac}(\mathbb{M})$ satisfying

$$\frac{\partial}{\partial t}\rho_t + \mathrm{div}\left(\rho_t(-\mathrm{grad}_\rho F)\right) = 0. \tag{A.2}$$

The equation (A.2) can be written as a weak form:

$$\left.\frac{d}{dt}\right|_{t=0} \int_{\mathbb{M}} \zeta\rho_t \, d\theta = -\int_{\mathbb{M}} \langle \nabla\zeta, \mathrm{grad}_\rho F\rangle \, \rho_t d\theta, \quad \forall \zeta \in C_c^\infty(\mathbb{M}).$$

We now verify that the equation (2.21) is the gradient flow in the above sense, of the potential V_k from (2.22). In our case the underlying Riemannian manifold $\mathbb{M}$ is $\mathbb{T}$, with the metric $d\theta^2$. Given ρ in $\mathbb{P}(\mathbb{T})$, recall the notation for J specified in (2.23). Then, we have

$$\mathrm{grad}_\rho V_k := -K\nabla(\sigma \cdot J). \tag{A.3}$$

where the inner product is the Euclidean one in $\mathbb{T} \subset \mathbb{R}^2$, and ∇ is the gradient of the Riemannian manifold $\mathbb{T}$; of course, $\nabla = \frac{d}{d\theta}$, but we keep the notation ∇ to be more consistent with the general formulation. In order to see (A.3), we note that for each one-parameter differentiable family, $t \mapsto g_t \in \mathbb{P}_{ac}(\mathbb{M})$ satisfying (A.1) with $g_0 = \rho$, the derivative of V_k given in (2.22) is

$$\begin{aligned} \left.\frac{d}{dt}\right|_{t=0} V_k(g_t) = & \quad -KJ \cdot \left.\frac{d}{dt}\right|_{t=0} \int_{\mathbb{T}} \sigma g_t \, d\theta = -K \left.\frac{d}{dt}\right|_{t=0} \int_{\mathbb{T}} (\sigma \cdot J) g_t \, d\theta \\ = & \quad -\int_{\mathbb{T}} \langle \nabla\varphi, K\nabla(\sigma \cdot J)\rangle \rho \, d\theta. \end{aligned}$$

This yields (A.3). On the other hand, note that since $J = Re^{i\phi}$,

$$\sigma(\theta) \cdot J = R\cos(\theta - \phi), \tag{A.4}$$

Therefore,

$$\mathrm{grad}_\rho V_k = -K\nabla(\sigma(\theta) \cdot J) = KR\sin(\theta - \phi).$$

We substitute this into (A.2) to get the gradient flow of V_k as the one-parameter family $t \in [0, \varepsilon) \mapsto \rho_t \in \mathbb{P}(\mathbb{T})$ satisfying

$$\partial_t \rho = \partial_\theta \left(\rho KR\sin(\theta - \phi)\right),$$

which is the same as (2.21), verifying the gradient flow structure. Moreover, this immediately implies

$$\left.\frac{d}{dt}\right|_{t=0} V_k(\rho_t) = -(KR)^2 \int_{\mathbb{T}} \sin(\theta - \phi)^2 \rho \, d\theta.$$

Of course, this expression can be directly obtained from the formula of V_k (2.24) and the definitions of the order parameters R and ϕ.

Additionally, using the Riemannian inner product on $(\mathbb{P}(\mathbb{T}), W_2)$, we have that the metric slope is given by

$$\int_{\mathbb{T}} K|\nabla(\sigma \cdot J)|^2 \rho \, d\theta = KR^2 \int_{\mathbb{T}} \sin^2(\theta - \phi)\rho \, d\theta.$$

A.2 The Hessian of the potential V_k

In this subsection, we explicitly compute the Hessian of the potential V_k via the Otto calculus. By direct calculations, the Hessian of V_k is given by

$$\langle \mathrm{Hess}_\rho \nabla\varphi, \nabla\varphi \rangle = -K \int_{\mathbb{T}} \nabla\varphi D^2(\sigma \cdot J)\nabla\varphi\rho(\theta) \, d\theta - K\left| \int_{\mathbb{T}} \nabla\varphi\rho \, d\theta \right|^2. \tag{A.5}$$

Using the Jensen inequality, this expression can be bounded from below by

$$-K\int_{\mathbb{T}}\nabla\varphi\big[D^2(\sigma\cdot J)+1\big]\nabla\varphi\rho(\theta)\;d\theta.$$

Consequently, the functional $V_k(\rho)$ is λ-convex with $\lambda=-2K$, as a functional on the formal Riemannian space $\mathbb{P}(\mathbb{T})$. In order to obtain (A.5), we proceed as follows. Let, $t\mapsto g_t\in\mathbb{P}(\mathbb{T})$, be the differentiable one-parameter family, with $g_0=\rho$, associated with $\varphi:\mathbb{T}\times\mathbb{R}\to\mathbb{R}$. Moreover, suppose

$$\begin{cases}\dfrac{d}{dt}\displaystyle\int_{\mathbb{T}}\zeta g_t\;d\theta=\int_{\mathbb{T}}\langle\nabla\zeta,\nabla\varphi\rangle g_t\;d\theta,\quad\forall\zeta\in C^\infty(\mathbb{T}),\\ \partial_t\varphi+\dfrac{\mid\nabla\varphi\mid^2}{2}=0.\end{cases}\qquad .\tag{A.6}$$

This family is a Riemannian geodesic in the formal Riemannian manifold $\mathbb{P}(\mathbb{T})$, in the sense of the Otto calculus. Then, we compute

$$\frac{d^2}{dt^2}\bigg|_{t=0}V_k(g_t)=\frac{d^2}{dt^2}\bigg|_{t=0}\left[K-\frac{K}{2}|J|^2\right]=-K\frac{d}{dt}\bigg|_{t=0}\int_{\mathbb{T}}\langle\nabla\varphi,\nabla(\sigma\cdot J)\rangle g_t\;d\theta.$$

Here, the last expression can be computed as

$$\begin{aligned}K\int_{\mathbb{T}}&\Big\langle\nabla\left(\frac{1}{2}\mid\nabla\varphi\mid^2\right),\nabla(\sigma\cdot J)\Big\rangle\rho\;d\theta-K\int_{\mathbb{T}}\Big\langle\nabla\varphi,\nabla\left[\sigma\cdot\frac{d}{dt}\bigg|_{t=0}J\right]\Big\rangle\rho\;d\theta\\&-K\int_{\mathbb{T}}\Big\langle\nabla\big\langle\nabla\varphi,\nabla\big(\sigma\cdot J\big)\big\rangle,\nabla\varphi\Big\rangle\rho\;d\theta\\&=:\mathcal{I}_{11}+\mathcal{I}_{12}+\mathcal{I}_{13}.\end{aligned}$$

where we used the system (A.6) in the first integral $\mathcal{I}_{11}$.

- (Estimate of $\mathcal{I}_{11} + \mathcal{I}_{13}$): We apply the identity

$$-\nabla\varphi\nabla^2\varphi\nabla\Psi + \langle\nabla\varphi, \nabla[\langle\nabla\varphi, \nabla\Psi\rangle]\rangle = \nabla\varphi D^2\Psi\nabla\varphi.$$

to obtain

$$\mathcal{I}_{11} + \mathcal{I}_{13} = -K\int_{\mathbb{T}} \nabla\varphi[D^2(\sigma\cdot J)]\nabla\varphi\rho\ d\theta. \tag{A.7}$$

- (Estimate of $\mathcal{I}_{12}$): In order to simplify $\mathcal{I}_{12}$ we use the identity

$$\begin{aligned} e\cdot\frac{d}{dt}\Big|_{t=0} J = \frac{d}{dt}\Big|_{t=0} e\cdot\int_{\mathbb{T}}\sigma g\ d\theta &= \int_{\mathbb{T}}(e\cdot\sigma)\partial_t g_t(\theta,0)\ d\theta \\ &= \int_{\mathbb{T}}\langle\nabla(e\cdot\sigma), \nabla\varphi\rangle\rho\ d\theta = e\cdot\int_{\mathbb{T}}\nabla\varphi\rho\ d\theta, \end{aligned}$$

which holds for any e in $\mathbb{R}^2$. In such an identity, we have used the fact that for any vector v in $\mathbb{R}^2$, $\nabla(v\cdot\sigma)$ is the orthogonal projection of v in the orthogonal subspace to $\sigma(\theta)$.

Consequently, by the same reason, $\mathcal{I}_{12}$ can be computed as

$$\mathcal{I}_{12} = -K\int_{\mathbb{T}}\langle\nabla\varphi, \nabla\left[\sigma\cdot\int_{\mathbb{T}}\nabla\varphi\rho\ d\theta_*\right]\rangle\rho\ d\theta = -K\left|\int_{\mathbb{T}}\nabla\varphi\rho\ d\theta\right|^2. \tag{A.8}$$

Finally, we combine (A.7) and (A.8) to obtain (A.5).

B Proof of Lemma 6.2

Below, we study the Lipschitz continuity of $R, \dot{R}$ and $\mathcal{M}(L_\gamma^+(t))$.

B.1 Lipschitz continuity of R

It follows from Lemma 5.2 and the facts

$$f = 0 \quad \text{for } \omega \notin [-M, M], \qquad R \le 1$$

that $\dot{R}$ is uniformly bounded:

$$|\dot{R}| \le M \iint_{\mathbb{T}\times[-M,M]} f(\theta,\omega,t)\ d\theta d\omega + KR \int_{\mathbb{T}} \rho(\theta,t)\ d\theta \le M + K. \tag{B.1}$$

This yields the Lipschitz continuity of R.

B.2 Lipschitz continuity of $\dot{R}$

: Note that the a priori condition $\min_{0\le t\le T} R(t) \ge \underline{R} > 0$ and the continuity of R yield that there exists a positive constant η such that

$$R(t) > \frac{\underline{R}}{2}, \quad t \in [0, T+\eta). \tag{B.2}$$

We next show that $\dot{R}$ is Lipschitz continuous on $[0, T+\eta)$. For this, it suffices to show that $\ddot{R}$ is uniformly bounded. Since

$$\frac{d^2}{dt^2}R^2 = 2(\dot{R})^2 + 2R\ddot{R}, \tag{B.3}$$

once we can show that $\frac{d^2}{dt^2}R^2$ is bounded, then, it follows from (B.1) and (B.2) that $\ddot{R}$ is bounded.

• Step A (uniform boundedness of $\frac{d^2}{dt^2}R^2$): We set

$$J(t) := \iint_{\mathbb{T}\times\mathbb{R}} e^{\mathrm{i}\theta^*} f(\theta^*,\omega,t)\ d\theta^* d\omega, \qquad \sigma(\theta) := e^{\mathrm{i}\theta}.$$

Then, we have

$$\sigma(\theta^*) \cdot \sigma(\theta) = \cos(\theta - \theta^*), \quad \sigma(\theta) \cdot J(t) = \iint_{\mathbb{T}\times\mathbb{R}} \cos(\theta - \theta^*) f \; d\theta^* d\omega.$$

This yields

$$\partial_\theta(\sigma(\theta) \cdot J) = -\iint_{\mathbb{T}\times\mathbb{R}} \sin(\theta - \theta^*) f \; d\theta^* d\omega. \tag{B.4}$$

Note that the K-S equation (2.15) can be written on $\mathbb{T} \times \mathbb{R}$ as

$$\partial_t f + \mathrm{div}_\theta(\omega\sigma^\perp f) + K\mathrm{div}_\theta(\partial_\theta(\sigma \cdot J) f) = 0. \tag{B.5}$$

Where div_θ and ∂_θ denote the divergence operator and gradient operator on $\mathbb{T} \subset \mathbb{R}^2$, endowed with angle metric, and for each σ in $\mathbb{T}$, $\sigma^\perp$ denotes the vector obtained by rotating σ by $\frac{\pi}{2}$ radians counterclockwise. We will denote the Laplacian on $\mathbb{T}$ by $\partial_{\theta\theta}$. We use (B.4), (B.5) and

$$R^2 = \iint_{\mathbb{T}\times\mathbb{R}} (\sigma \cdot J) f(\theta, \omega, t) \; d\theta d\omega,$$

to obtain

$$\frac{dR^2}{dt} = 2\Big[\iint_{\mathbb{T}\times\mathbb{R}} \partial_\theta(\sigma \cdot J) \cdot \omega\sigma^\perp f \; d\theta d\omega + K\iint_{\mathbb{T}\times\mathbb{R}} (\partial_\theta(\sigma \cdot J))^2 f \; d\theta d\omega\Big]. \tag{B.6}$$

Then, we claim:

$$
\begin{aligned}
\frac{d^2}{dt^2}R^2 = 4K^2 &\iint_{\mathbb{T}\times\mathbb{R}} \partial_\theta(\sigma\cdot J)\partial_{\theta\theta}(\sigma\cdot J)\partial_\theta(\sigma\cdot J)\ f\ d\theta d\omega \\
&+ 4K \iint_{\mathbb{T}\times\mathbb{R}} \omega\sigma^\perp\cdot(\partial_{\theta\theta}\sigma\cdot J)\partial_\theta(\sigma\cdot J) f\ d\theta d\omega \\
&+ 4K^2\left|\iint_{\mathbb{T}\times\mathbb{R}} \partial_\theta(\sigma\cdot J) f\ d\theta d\omega\right|^2 \\
&+ 4K \iint_{\mathbb{T}\times\mathbb{R}} \partial_\theta(\sigma\cdot J) f\ d\theta d\omega \cdot \iint_{\mathbb{T}\times\mathbb{R}} \omega\sigma^\perp f\ d\theta d\omega \\
&+ 2K \iint_{\mathbb{T}\times\mathbb{R}} \partial_\theta(\sigma\cdot J)\partial_{\theta\theta}(\sigma\cdot J)\omega\sigma^\perp f\ d\theta d\omega \\
&+ 2 \iint_{\mathbb{T}\times\mathbb{R}} \omega\sigma^\perp\partial_{\theta\theta}(\sigma\cdot J)\omega\sigma^\perp f\ d\theta d\omega \\
&+ 2K \iint_{\mathbb{T}\times\mathbb{R}} \omega\sigma^\perp f\ d\theta d\omega \cdot \iint_{\mathbb{T}\times\mathbb{R}} \partial_\theta(\sigma\cdot J) f\ d\theta d\omega \\
&+ 2\left|\iint_{\mathbb{T}\times\mathbb{R}} \omega\sigma^\perp f\ d\theta d\omega\right|^2.
\end{aligned}
\tag{B.7}
$$

Proof of claim (B.7): It follows from (B.6) that we have

$$
\frac{d^2}{dt^2}R^2 = 2\frac{d}{dt}\left[\iint_{\mathbb{T}\times\mathbb{R}} \partial_\theta(\sigma\cdot J)\omega\sigma^\perp f\ d\theta d\omega + K\iint_{\mathbb{T}\times\mathbb{R}} (\partial_\theta(\sigma\cdot J))^2 f\ d\theta d\omega\right]. \tag{B.8}
$$

• (The second integral in the R.H.S. of (B.8)): By direct calculations, we have

$$
\begin{aligned}
&\frac{d}{dt}\iint_{\mathbb{T}\times\mathbb{R}} |\partial_\theta(\sigma\cdot J)|^2 f\ d\theta d\omega \\
&\qquad = K\iint_{\mathbb{T}\times\mathbb{R}} \partial_\theta|\partial_\theta(\sigma\cdot J)|^2\partial_\theta(\sigma\cdot J) f\ d\theta d\omega \\
&+ \iint_{\mathbb{T}\times\mathbb{R}} \omega\sigma^\perp\partial_\theta|\partial_\theta(\sigma\cdot J)|^2 f\ d\theta d\omega \\
&\qquad + 2\iint_{\mathbb{T}\times\mathbb{R}} \partial_\theta(\sigma\cdot J)\partial_\theta(\sigma\cdot\partial_t J) f\ d\theta d\omega \\
&\qquad = 2K\iint_{\mathbb{T}\times\mathbb{R}} \partial_\theta(\sigma\cdot J)\partial_{\theta\theta}(\sigma\cdot J)\partial_\theta(\sigma\cdot J) f\ d\theta d\omega \\
&\qquad + 2\iint_{\mathbb{T}\times\mathbb{R}} \omega\sigma^\perp\partial_{\theta\theta}(\sigma\cdot J)\partial_\theta(\sigma\cdot J) f\ d\theta d\omega \\
&\qquad + 2K\iint_{\mathbb{T}\times\mathbb{R}} \partial_\theta(\sigma\cdot J)\partial_\theta\Big[\sigma\cdot\big(\iint_{\mathbb{T}\times\mathbb{R}}\partial_\theta(\sigma\cdot J) f d\theta d\omega\big)\Big] f\ d\theta d\omega \\
&\qquad + 2\iint_{\mathbb{T}\times\mathbb{R}} \partial_\theta(\sigma\cdot J)\partial_\theta\Big[\sigma\cdot\big(\iint_{\mathbb{T}\times\mathbb{R}}\omega\sigma^\perp f\ d\theta d\omega\big)\Big] f\ d\theta d\omega,
\end{aligned}
\tag{B.9}
$$

where in the second equality, we used the identity:

$$
\begin{aligned}
e\cdot\partial_t J &= \frac{d}{dt} e\cdot\iint_{\mathbb{T}\times\mathbb{R}} \sigma f\ d\theta d\omega = \iint_{\mathbb{T}\times\mathbb{R}} (e\cdot\sigma)\partial_t f\ d\theta d\omega \\
&= K\iint_{\mathbb{T}\times\mathbb{R}} \partial_\theta(e\cdot\sigma)\partial_\theta(\sigma\cdot J) f\ d\theta d\omega + \iint_{\mathbb{T}\times\mathbb{R}} \omega\sigma^\perp\partial_\theta(e\cdot\sigma) f\ d\theta d\omega \\
&= Ke\cdot\iint_{\mathbb{T}\times\mathbb{R}} \partial_\theta(\sigma\cdot J) f\ d\theta d\omega + e\cdot\iint_{\mathbb{T}\times\mathbb{R}} \omega\sigma^\perp f\ d\theta d\omega,
\end{aligned}
$$

which holds for any e in $\mathbb{R}^2$. In the above identity, we have used the fact that for any vector v in $\mathbb{R}^2$, $\partial_\theta(v\cdot\sigma)$ is the orthogonal projection of v in the orthogonal subspace to $\sigma(\theta)$, and the fact that, by definition, $\omega\sigma^\perp(\theta)$ is contained in the same subspace as well. By similar arguments, we also get

$$
\begin{aligned}
\frac{d}{dt} &\iint_{\mathbb{T}\times\mathbb{R}} (\partial_\theta(\sigma \cdot J))^2 f \ d\theta d\omega \\
&\quad = 2K \iint_{\mathbb{T}\times\mathbb{R}} \partial_\theta(\sigma \cdot J_f)\partial_{\theta\theta}(\sigma \cdot J)\partial_\theta(\sigma \cdot J) f \ d\theta d\omega \\
&\quad + 2 \iint_{\mathbb{T}\times\mathbb{R}} \omega\sigma^\perp \partial_{\theta\theta}(\sigma \cdot J)\partial_\theta(\sigma \cdot J) f \ d\theta d\omega \\
&\quad + 2K \left| \iint_{\mathbb{T}\times\mathbb{R}} \partial_\theta(\sigma \cdot J) f \ d\theta d\omega \right|^2 \\
&+ 2 \iint_{\mathbb{T}\times\mathbb{R}} \partial_\theta(\sigma \cdot J) f \ d\theta d\omega \cdot \iint_{\mathbb{T}\times\mathbb{R}} \omega\sigma^\perp f \ d\theta d\omega.
\end{aligned}
\tag{B.10}
$$

• (The first integral in the R.H.S. of (B.8)): By direct calculation, we have

$$
\begin{aligned}
&\frac{d}{dt}\iint_{\mathbb{T}\times\mathbb{R}} \partial_\theta(\sigma\cdot J)\omega f\ d\theta d\omega \\
&\qquad = K\iint_{\mathbb{T}\times\mathbb{R}} \partial_\theta(\sigma\cdot J)\partial_\theta[\partial_\theta(\sigma\cdot J)\cdot\omega\sigma^\perp]f\ d\theta d\omega \\
&\qquad + \iint_{\mathbb{T}\times\mathbb{R}} \omega\sigma^\perp\cdot\partial_\theta[\partial_\theta(\sigma\cdot J)\omega\sigma^\perp]f\ d\theta d\omega \\
&+ \iint_{\mathbb{T}\times\mathbb{R}} \partial_\theta(\sigma\cdot\partial_t J)\cdot\omega\sigma^\perp f\ d\theta d\omega \\
&\qquad = K\iint_{\mathbb{T}\times\mathbb{R}} \partial_\theta(\sigma\cdot J)\partial_{\theta\theta}(\sigma\cdot J)\omega\sigma^\perp f\ d\theta d\omega \\
&+ \iint_{\mathbb{T}\times\mathbb{R}} \omega\sigma^\perp\cdot\partial_{\theta\theta}(\sigma\cdot J)\omega\sigma^\perp f\ d\theta d\omega \\
&\qquad + K\iint_{\mathbb{T}\times\mathbb{R}} \partial_\theta[\sigma\cdot\iint_{\mathbb{T}\times\mathbb{R}} \partial_\theta(\sigma\cdot J)f\ d\theta d\omega]\omega\sigma^\perp f\ d\theta d\omega \\
&\qquad + \iint_{\mathbb{T}\times\mathbb{R}} \partial_\theta[\sigma\cdot\iint_{\mathbb{T}\times\mathbb{R}} \omega\sigma^\perp f\ d\theta d\omega]\omega\sigma^\perp f\ d\theta d\omega \\
&\qquad = K\iint_{\mathbb{T}\times\mathbb{R}} \partial_\theta(\sigma\cdot J)\partial_{\theta\theta}(\sigma\cdot J)\omega\sigma^\perp f\ d\theta d\omega \\
&\qquad + \iint_{\mathbb{T}\times\mathbb{R}} \omega\sigma^\perp\cdot\partial_{\theta\theta}(\sigma\cdot J)\omega\sigma^\perp f\ d\theta d\omega \\
&\qquad + K\iint_{\mathbb{T}\times\mathbb{R}} \partial_\theta(\sigma\cdot J)f\ d\theta d\omega\cdot \\
&\iint_{\mathbb{T}\times\mathbb{R}} \omega\sigma^\perp f\ d\theta d\omega + \left|\iint_{\mathbb{T}\times\mathbb{R}} \omega\sigma^\perp f\ d\theta d\omega\right|^2,
\end{aligned}
\tag{B.11}
$$

where in the second and third equalities above, we have used the same tools we used to obtain (B.10). Finally, in (B.8), we combine (B.10) and (B.11) to prove the claim.

• Step B (uniform boundedness of $\ddot{R}$): It follows from the relation (B.7) that

$$\left|\frac{d^2}{dt^2}R^2\right| \le 4(K^2+KM)+2KM+2M^2, \tag{B.12}$$

Then, we use the relations (B.2) and (B.3) to see that for $t \in [0, T+\eta)$,

$$\begin{aligned}\underline{R}|\ddot{R}| \le 2|R\ddot{R}| &\le \left|\frac{d^2}{dt^2}R^2\right| + 2|\dot{R}|^2 \\ &\le 4(K^2+KM)+2KM+2M^2+2(M+K)^2.\end{aligned}$$

This yields the desired bound estimate for $\ddot{R}$:

$$|\ddot{R}| \le \frac{1}{\underline{R}}\Big[4(K^2+KM)+2KM+2M^2+2(M+K)^2\Big], \quad t \in [0, T+\eta).$$

and implies the Lipschitz continuity of $\dot{R}$ on $[0, T+\eta)$.

B.3 Lipschitz continuity of $\mathcal{M}(L_\gamma^+(t))$

It follows from (5.27) that we have

$$\begin{aligned}\frac{d}{dt}\mathcal{M}(L_\gamma^+(t)) = KR\cos\gamma \int_{-M}^{M} \big[B_{-,\omega}(t)+B_{+,\omega}(t)\big]\, d\omega \\ + \int_{-M}^{M} \big(\dot{\phi}(t)-\omega\big)\big[B_{+,\omega}(t)-B_{-,\omega}(t)\big]\, d\omega. \quad \text{(B.13)}\end{aligned}$$

Hence, we use (5.17), $R \le 1$, Lemma 5.3, (B.2) and

$$\int_{-M}^{M} \big[B_{-,\omega}(t)+B_{+,\omega}(t)\big]\, d\omega \le 2M\|f_0\|_{L^\infty(\mathbb{T}\times[-M,M])}e^{K(T+\eta)}$$

to obtain

$$\begin{aligned}\left|\frac{d}{dt}\mathcal{M}(L_\gamma^+(t))\right| &\le (KR+|\dot{\phi}|+M)\int_{-M}^{M} \big[B_{-,\omega}(t)+B_{+,\omega}(t)\big]\, d\omega \\ &\le 2M\left[K+\frac{2M}{\underline{R}}+K\left(1-\frac{\underline{R}}{2}\right)+M\right]\|f_0\|_{L^\infty(\mathbb{T}\times[-M,M])}e^{K(T+\eta)},\end{aligned}$$

on $[0, T+\eta)$. This concludes the proof of Lemma 6.2.

C Proof of Lemma 6.5

Suppose that f_0, R and K satisfy

$$\|f_0\|_{L^\infty(\mathbb{T}\times[-M,M])} < \infty, \quad \min_{0\le t\le T} R(t) > \underline{R}, \quad \dot{R}(T) = K\mu$$
$$\text{and} \quad K^2\mu > \frac{M^2}{2\underline{R}^2} - \frac{3M^2}{4\underline{R}} \tag{C.1}$$

for some $T \in (0,\infty)$ and some positive constants $\underline{R}$, and μ. Then, we claim: there exist positive constants $d := d(K, M, \underline{R}, \mu)$ and $E_3 := E_3(K, M, \underline{R}, \mu)$ satisfying

$$\dot{R} > 0 \quad \text{in} \quad [T, T+d), \qquad R(T+d) - R(T) \ge \frac{1}{12}\underline{R}\mu - E_3.$$

Note that Lemma 5.2 yields

$$\dot{R} = -\iint_{\mathbb{T}\times\mathbb{R}} \sin(\theta-\phi)\omega f(\theta,\omega,t)\ d\theta d\omega + KR\int_{\mathbb{T}} \sin^2(\theta-\phi)\rho(\theta,t)\ d\theta.$$

We define a function $S\ :\ [0,\infty) \to \mathbb{R}$ given by

$$S(t) = KR(t)\iint_{\mathbb{T}\times\mathbb{R}} \sin^2(\theta-\phi) f(\theta,\omega,t)\ d\theta d\omega. \tag{C.2}$$

We remark that for identical oscillator case, this coincides with metric slope.

• Step A (Derivation of differential inequalities for R): For any $t \ge T$ satisfying $R(t) \ge \underline{R}$, we claim:

$$-\frac{M^2}{2K\underline{R}} + \frac{1}{2}S(t) \le \dot{R}(t) \le \frac{3}{2}S(t) + \frac{M^2}{2\underline{R}K}. \tag{C.3}$$

Proof of claim (C.3): Suppose that

$$R(t) \geq \underline{R}.$$

Then, it follows from Lemma 5.2 that we have

$$\begin{aligned} \dot{R}(t) = &- \iint_{\mathbb{T}\times\mathbb{R}} \sin(\theta-\phi)\omega f(\theta,\omega,t)\ d\theta d\omega \\ &+ KR(t) \iint_{\mathbb{T}\times\mathbb{R}} \sin^2(\theta-\phi) f(\theta,\omega,t)\ d\theta d\omega. \end{aligned} \tag{C.4}$$

For the first term in (C.4), we use Young's inequality:

$$|ab| \leq \frac{a^2}{2\varepsilon} + \frac{b^2}{2}\varepsilon, \quad \text{with} \ \ \varepsilon = KR$$

and the relations $|\omega| \leq M, \quad R(t) \geq \underline{R}$ to see

$$\begin{aligned} &\Big| - \iint_{\mathbb{T}\times\mathbb{R}} \sin(\theta-\phi)\omega f(\theta,\omega,t)\ d\theta d\omega \Big| \\ &\quad \leq \frac{1}{2KR(t)} \iint_{\mathbb{T}\times\mathbb{R}} \omega^2 f\ d\theta d\omega \\ &+ \frac{KR(t)}{2} \int_{\mathbb{T}\times\mathbb{R}} \sin^2(\theta-\phi) f(\theta,\omega,t)\ d\theta d\omega \leq \frac{M^2}{2K\underline{R}} + \frac{1}{2}S(t). \end{aligned} \tag{C.5}$$

We combine (C.4) and (C.5) to verify the claim (C.3).

Note that the relations (C.2) and (A.4) imply

$$\begin{aligned} R(t)S(t) &= KR^2(t) \iint_{\mathbb{T}\times\mathbb{R}} \sin^2(\theta-\phi) f(\theta,\omega,t)\ d\theta d\omega \\ &= K \iint_{\mathbb{T}\times\mathbb{R}} (-R\sin(\theta-\phi))^2 f(\theta,\omega,t)\ d\theta d\omega \\ &= K \iint_{\mathbb{T}\times\mathbb{R}} |\partial_\theta(\sigma\cdot J)|^2 f\ d\theta d\omega. \end{aligned} \tag{C.6}$$

Then, we use (C.6) and (B.10) to obtain

$$\begin{aligned}\frac{d}{dt}(R(t)S(t)) = 2K^2 &\iint_{\mathbb{T}\times\mathbb{R}} \partial_\theta(\sigma\cdot J)\partial_{\theta\theta}(\sigma\cdot J)\partial_\theta(\sigma\cdot J) f \ d\theta d\omega \\ &+ 2K \iint_{\mathbb{T}\times\mathbb{R}} \omega\sigma^\perp \partial_{\theta\theta}(\sigma\cdot J)\partial_\theta(\sigma\cdot J) f \ d\theta d\omega \\ &+ 2K^2 \left| \iint_{\mathbb{T}\times\mathbb{R}} \partial_\theta(\sigma\cdot J) f \ d\theta d\omega \right|^2 \\ &+ 2K \iint_{\mathbb{T}\times\mathbb{R}} \partial_\theta(\sigma\cdot J) f d\theta d\omega \iint_{\mathbb{T}\times\mathbb{R}} \omega\sigma^\perp f \ d\theta d\omega.\end{aligned} \tag{C.7}$$

By Young's inequality, we get

$$\begin{aligned}\frac{d}{dt}(R(t)S(t)) &\geq -2K^2 \iint_{\mathbb{T}\times\mathbb{R}} |\partial_\theta(\sigma\cdot J) f|^2 f \ d\theta d\omega \\ &- K^2 \iint_{\mathbb{T}\times\mathbb{R}} |\partial_\theta(\sigma\cdot J) f|^2 f \ d\theta d\omega \\ &- \iint_{\mathbb{T}\times\mathbb{R}} \omega^2 f \ d\theta d\omega - K^2 \iint_{\mathbb{T}\times\mathbb{R}} |\partial_\theta(\sigma\cdot J)|^2 f \ d\theta d\omega \\ &- \iint_{\mathbb{T}\times\mathbb{R}} \omega^2 f \ d\theta d\omega \\ &\geq -4KR(t)S(t) - 2M^2.\end{aligned}$$

Then, Grownwall's lemma yields

$$R(t)S(t) \geq \left(R(T)S(T) + \frac{M^2}{2K}\right)e^{-4K(t-T)} - \frac{M^2}{2K}.$$

Since $S \geq 0$ and $R \leq 1$, we also obtain

$$S(t) \geq R(t)S(t) \geq \left(R(T)S(T) + \frac{M^2}{2K}\right)e^{-4K(t-T)} - \frac{M^2}{2K}. \tag{C.8}$$

- Step B (Lower bound of R): We next claim: for some $d > 0$,

$$R \geq R(T), \quad \text{in} \ \ [T, T+d). \tag{C.9}$$

For the proof of claim (C.9), we first define a positive constant d by the following implicit relation:

$$\left[\frac{1}{3}\underline{R}\dot{R}(T) - \left(\frac{M^2}{6\underline{R}K} - \frac{M^2}{4K}\right)\right]e^{-4Kd} - \frac{1}{4}\frac{M^2}{K} - \frac{M^2}{2\underline{R}K} = 0. \tag{C.10}$$

The unique existence of such d is guaranteed by the condition (C.1). We introduce a set $\mathfrak{T}$ and its supremum as follows.

$$\mathfrak{T} := \Big\{t \in [T, T+d) \ : \ R(t^*) \geq R(T) \ \ \forall t^* \in [T, t]\Big\}, \quad \tau := \sup \mathfrak{T}.$$

Since $T \in \mathfrak{T}$, the set $\mathfrak{T}$ is non-empty and τ is well defined. To prove a claim (C.9), it suffices to show

$$\tau \geq T + d.$$

Suppose not, i.e. $\tau < T + d$. By the continuity of R which is guaranteed by Lemma 6.1 and definition of τ, we have

$$R(\tau) = R(T), \quad \dot{R}(\tau) \leq 0.$$

On the other hand, definition of τ allows us to use inequality (C.3) in (C.8), for every t in the interval $[T, \tau]$. By doing so, we obtain

$$\begin{aligned} S(t) &\geq \left(\frac{2}{3}R(T)\dot{R}(T) - R(T)\frac{M^2}{3\underline{R}K} + \frac{M^2}{2K}\right)e^{-4K(t-T)} - \frac{M^2}{2K} \\ &\geq \left(\frac{2}{3}R(T)\dot{R}(T) - \frac{M^2}{3\underline{R}K} + \frac{M^2}{2K}\right)e^{-4K(t-T)} - \frac{M^2}{2K}, \quad \text{in } [T, \tau), \end{aligned}$$

where we have used the fact that $R \leq 1$. Hence, another application of (C.3) yields

$$
\begin{aligned}
\dot{R}(\tau) &\geq \left[\frac{1}{3}\underline{R}\dot{R}(T) - \left(\frac{M^2}{6\underline{R}K} - \frac{M^2}{4K}\right)\right] e^{-4K(\tau - T)} - \frac{1}{4}\frac{M^2}{K} - \frac{M^2}{2\underline{R}K}. \\
&> \left[\frac{1}{3}\underline{R}\dot{R}(T) - \left(\frac{M^2}{6\underline{R}K} - \frac{M^2}{4K}\right)\right] e^{-4Kd} - \frac{1}{4}\frac{M^2}{K} - \frac{M^2}{2\underline{R}K} \\
&= 0,
\end{aligned}
$$

In the second inequality, we have used the condition on K in (C.1), the assumption $\tau < T + d$, and the strict monotonicty of the exponential function. Thus, we reach a contradiction. Hence, we conclude $\tau = T+d$. By the previous argument, we have

$$
\dot{R}(t) \geq \left[\frac{1}{3}\underline{R}\dot{R}(T) - \left(\frac{M^2}{6\underline{R}K} - \frac{M^2}{4K}\right)\right] e^{-4K(t-T)} - \frac{1}{4}\frac{M^2}{K} - \frac{M^2}{2\underline{R}K}, \quad \text{in } [T, T+d). \tag{C.11}
$$

On the other hand, we use definition of d to see

$$
d = \frac{1}{4K}\log\left[\frac{1}{3}\underline{R}\dot{R}(T) - \left(\frac{M^2}{6\underline{R}K} - \frac{M^2}{4K}\right)\right] - \frac{1}{4K}\log\left(\frac{1}{4}\frac{M^2}{K} + \frac{M^2}{2\underline{R}K}\right).
$$

For notational simplicity, we set

$$
a := \frac{1}{3}\underline{R}\dot{R}(T) - \left(\frac{M^2}{6\underline{R}K} - \frac{M^2}{4K}\right) \quad \text{and} \quad b := \frac{1}{4}\frac{M^2}{K} + \frac{M^2}{2\underline{R}K}.
$$

It follows from (C.11) that we have

$$
\begin{aligned}
R(T+d) - R(d) &= \int_T^{T+d} \frac{d}{dt}R(t)\ dt \geq \int_T^{T+d} [ae^{-4K(t-T)} - b]\ dt, \\
&= [-\frac{a}{4K}e^{-4K(t-T)} - bt]\Big|_{t=T}^{t=T+d} \\
&= \frac{a}{4K}\left(1 - e^{-4Kd}\right) - bd.
\end{aligned}
$$

Then, by (C.10), the assumption that $\dot{R}(T) = K\mu$ and definition of a and b, we have the desired result.

$$
\begin{aligned}
&R(T+d) - R(d) \\
&\qquad \geq \left[\frac{1}{12}\underline{R}\frac{\dot{R}(T)}{K} - \frac{1}{4K}\left(\frac{M^2}{6\underline{R}K} - \frac{M^2}{4K}\right)\right]\left[1 - \frac{\frac{M^2}{4K} + \frac{M^2}{2\underline{R}K}}{\frac{1}{3}\underline{R}\dot{R}(T) - \left(\frac{M^2}{6\underline{R}K} - \frac{M^2}{4K}\right)}\right] \\
&\qquad - \left(\frac{M^2}{4K} + \frac{M^2}{2\underline{R}K}\right)\left[\frac{1}{4K}\log\frac{\frac{1}{3}\underline{R}\dot{R}(T) - \left(\frac{M^2}{6\underline{R}K} - \frac{M^2}{4K}\right)}{\frac{M^2}{4K} + \frac{M^2}{2\underline{R}K}}\right] \\
&\qquad \geq \frac{1}{12}\underline{R}\mu - \frac{1}{12}\underline{R}\mu\frac{\frac{M^2}{4K} + \frac{M^2}{2\underline{R}K}}{\frac{1}{3}\underline{R}K\mu - \left(\frac{M^2}{6\underline{R}K} - \frac{M^2}{4K}\right)} \\
&\qquad - \frac{1}{4K}\left(\frac{M^2}{6\underline{R}K} - \frac{M^2}{4K}\right)\left[1 - \frac{\frac{M^2}{4K} + \frac{M^2}{2\underline{R}K}}{\frac{1}{3}\underline{R}K\mu - \left(\frac{M^2}{6\underline{R}K} - \frac{M^2}{4K}\right)}\right] \\
&\qquad - \left(\frac{M^2}{4K} + \frac{M^2}{2\underline{R}K}\right)\left[\frac{1}{4K}\log\frac{\frac{1}{3}\underline{R}K\mu - \left(\frac{M^2}{6\underline{R}K} - \frac{M^2}{4K}\right)}{\frac{M^2}{4K} + \frac{M^2}{2\underline{R}K}}\right].
\end{aligned}
$$

D Proof of Proposition 6.1

Suppose that the assumptions ($\mathcal{H}1$) - ($\mathcal{H}3$) hold, and assume that there exists $t_0 \geq 0$ and $\underline{\mathrm{R}} > 0$ such that

$$
R(t_0) \geq R_0, \quad \inf_{0 \leq t \leq t_0} R(t) > \underline{\mathrm{R}}.
$$

Let $t \geq t_0$ be an instant satisfying $\dot{R}(t) \leq 0$. Then for such t, in which R is in non-increasing mode, we claim:

$$\mathcal{M}(L^+_{\frac{\pi}{3}}(t)) \geq \frac{1}{2}(R(t_0)+1) - E_1. \tag{D.1}$$

For the proof of claim, we consider a set $\mathcal{N}(t_0)$ consisting of non-increasing moments of R after t_0:

$$\mathcal{N}(t_0) := \Big\{ t \geq t_0 \ : \ \dot{R}(t) \leq 0 \Big\},$$

and the set $\mathcal{T}(t_0)$:

$$\mathcal{T}(t_0) := \Big\{ s \in [t_0, \infty) \ : \ \text{(D.1)} \ \text{ holds } \ \ \forall t \in [t_0, s] \cap \mathcal{N}(t_0) \Big\}.$$

We set

$$T^*(t_0) := \sup \mathcal{T}(t_0).$$

Notice that $[t_0, T^*(t_0)) \subset \mathcal{T}(t_0)$ and it suffices now to prove $T^*(t_0) = \infty$. Since the proof is rather long, we split its proof into several steps.

• Step A (the set $\mathcal{T}(t_0)$ is not empty):

If $\dot{R}(t_0) > 0$, the defining relation of the set $\mathcal{T}(t_0)$ holds trivially. Thus, $t_0 \in \mathcal{T}(t_0)$.
If $\dot{R}(t_0) \leq 0$, then it follows from Lemma 6.3 that (D.1) holds for $t = t_0$.

Thus, $t_0 \in \mathcal{T}(t_0)$. In any case, the set $\mathcal{T}(t_0)$ is not empty. Thus, its supremum exists and lies in the set $[t_0, \infty]$.

• Step B (the supremum $T^*(t_0) = \infty$): Suppose not, i.e.,

$$T^*(t_0) < \infty.$$

⋄ Step B.1: We want to show

$$R(t) \geq R(t_0) - 2E_1 - E_2, \quad t \in [t_0, T^*(t_0)), \tag{D.2}$$

where E_1 and E_2 were defined in assumption $(\mathcal{H}3)$. To see this, first note that

$$R(t_0) \geq R(t_0) - 2E_1 - E_2$$

from $(\mathcal{H}3)$. Second, R and $\dot{R}$ are Lipschitz continuous due to the Lemma 6.2. Thus, for each $t \in [t_0, T^*(t_0))$ such that $\dot{R}(t) \leq 0$, from the definition of $T^*(t_0)$, we have (D.1), which implies from Lemma 6.3

$$R(t) \geq R(t_0) - 2E_2 - E_1.$$

This means the quantity $R(t_0) - 2E_2 - E_1$ is a lower bound for R in $[t_0, T^*(t_0))$. This shows the claim.

⋄ Step B.2: we claim:

$$\dot{R}(T^*(t_0)) = 0. \tag{D.3}$$

This property comes directly form the continuity of $\dot{R}$ and $\mathcal{M}(L^+_{\frac{\pi}{3}}(t))$.

♣ Case A: Suppose $\dot{R}(T^*(t_0)) > 0$. Then there exists a time interval $(T^*(t_0) - \eta, T^*(t_0) + \eta)$ such that

$$\dot{R}(t) > 0 \quad \text{for} \quad t \in (T^*(t_0) - \eta, T^*(t_0) + \eta),$$

which contradicts to definition of $T^*(t_0) = \sup \mathcal{T}$.

♣ Case B: Suppose $\dot{R}(T^*(t_0)) < 0$. In this case, we have

$$\dot{R}(t) < 0 \quad \text{for} \quad t \in (T^*(t_0) - \eta, T^*(t_0) + \eta).$$

By Lemma 6.4, we have

$$\frac{d}{dt}\mathcal{M}(L^+_{\frac{\pi}{3}}(t)) \geq 0 \quad \text{for} \;\; t \in (T^*(t_0) - \eta, T^*(t_0) + \eta).$$

Here we used Step B.1 to satisfy the condition $R(t) \geq \underline{R}$. This gives

$$\mathcal{M}(L^+_{\frac{\pi}{3}}(t)) \geq \mathcal{M}(L^+_{\frac{\pi}{3}}(T^*(t_0))) \geq \frac{1}{2}(R(t_0)+1)-E_1 \quad \text{for} \quad T^*(t_0) \leq t < T^*(t_0)+\eta,$$

which also contradicts to definition of $T^*(t_0)$. Thus, we obtain the desired result (D.3).

⋄ Step B.3: In this part, we want to show that the mass in the interval $L^+_{\frac{\pi}{3}}$ at $T^*(t_0)$ satisfies

$$\mathcal{M}(L^+_{\frac{\pi}{3}}(T^*(t_0))) \geq \frac{1}{2}(R(t_0) + 1) - E_1.$$

Notice that from Step B.2 and Lemma 6.3, we have

$$\mathcal{M}(L^+_{\frac{\pi}{3}}(T^*(t_0))) \geq \frac{1}{2}\Big(R(T^*(t_0)) + 1\Big) - E_1$$

thus it suffices to show $R(T^*(t_0)) \geq R(t_0)$. Now, consider the two cases:

♣ Case A ($\mathcal{N}(t_0) \cap [t_0, T^*(t_0)) = \emptyset$): In this case, since we have $\dot{R}(t) \geq 0$ for $t \in [t_0, T^*(t_0)]$, we get

$$R(T^*(t_0)) \geq R(t_0).$$

♣ Case B ($\mathcal{N}(t_0) \cap [t_0, T^*(t_0)) \neq \emptyset$): We define $t_s := \sup\Big(\mathcal{N}(t_0) \cap [t_0, T^*(t_0))\Big)$.

- Suppose there exists a sequence $\{t_k\} \subset \mathcal{N}(t_0) \cap [t_0, T^*(t_0))$ such that $t_k \uparrow T^*(t_0)$ and $\dot{R}(t_k) \leq 0$ for all $k \in \mathbb{N}$, i.e., $t_s = T^*(t_0)$. Then, we have

  $$\mathcal{M}(L^+_{\frac{\pi}{3}}(t_k)) \geq \frac{1}{2}(R(t_0) + 1) - E_1 \quad \text{for} \quad k \in \mathbb{N}.$$

  Thus, by the continuity of $\mathcal{M}(L^+_{\frac{\pi}{3}}(t))$, we obtain

  $$\mathcal{M}(L^+_{\frac{\pi}{3}}(T^*(t_0))) \geq \frac{1}{2}(R(t_0) + 1) - E_1.$$

- For the case of $t_s < T^*(t_0)$, we have

  $$\dot{R}(t) > 0 \quad \text{for} \ \ t \in (t_s, T^*(t_0)).$$

  Lemma 6.3 and $\dot{R}(t_s) \leq 0$ imply

  $$\mathcal{M}(L^+_{\frac{\pi}{3}}(t_s)) \geq \frac{1}{2}(R(t_0) + 1) - E_1.$$

  We now investigate the mass $\mathcal{M}(L^+_{\frac{\pi}{3}}(t))$ for $t \in (t_s, T^*(t_0)]$:

  1. If $\dot{R}(t) < K\mu$ for all $t \in (t_s, T^*(t_0)]$, by Lemma 6.4, the mass is non-decreasing, i.e., $\frac{d}{dt}\mathcal{M}(L^+_{\frac{\pi}{3}}(t)) \geq 0$ for $t \in (t_s, T^*(t_0)$. Thus, we attain

     $$\mathcal{M}(L^+_{\frac{\pi}{3}}(T^*(t_0))) \geq \mathcal{M}(L^+_{\frac{\pi}{3}}(t_s)) \geq \frac{1}{2}(R(t_0) + 1) - E_1.$$

2. Suppose $\dot{R}(t) \geq K\mu$ for some $t \in (t_s, T^*(t_0)]$. Since $\dot{R}(t_s) \leq 0$ and the continuity of $\dot{R}$, there exists t_c such that

$$\dot{R}(t) < K\mu \quad \text{for} \quad t \in (t_s, t_c) \quad \text{and} \quad \dot{R}(t_c) = K\mu.$$

By Lemma 6.5, there exists a positive constant d such that

$$\dot{R}(t) > 0 \quad \text{for} \quad t \in [t_c, t_c + d) \quad \text{and} \quad R(t_c + d) - R(t_c) \geq \frac{R_0}{24}\mu - E_3.$$

Since we have $\dot{R}(T^*(t_0)) = 0$ from Step B.2 and the contiunity of $\dot{R}$, we attain $t_c + d < T^*(t_0)$. By the definition of t_s, we have $\dot{R}(t) > 0$ for $t \in (t_s, T^*(t_0))$, which implies

$$R(T^*(t_0)) - R(t_s) \geq R(t_c + d) - R(t_c) \geq \frac{R_0}{24}\mu - E_3$$

By Step B.1 and ($\mathcal{H}3$), we obtain

$$\begin{aligned} R(T^*(t_0)) &\geq R(t_s) + \frac{R_0}{24}\mu - E_3 \\ &\geq R(t_0) + \frac{R_0}{24}\mu - E_3 - E_2 - 2E_1 \\ &\geq R(t_0). \end{aligned}$$

We again use Lemma 6.3 for the result of Step B.2 to get

$$\mathcal{M}(L^+_{\frac{\pi}{3}}(T^*(t_0))) \geq \frac{1}{2}(R(T^*(t_0)) + 1) - E_1 \geq \frac{1}{2}(R(t_0) + 1) - E_1,$$

which conclude Step B.3.

$\diamond$ Step B.4: Finally, we show $T^*(t_0) = \infty$. Since $\dot{R}(T^*(t_0)) = 0$ from Step B.2 and the continuity of $\dot{R}$, there is a small time interval such that

$$\dot{R}(t) < K\mu \quad \text{for} \quad t \in (T^*(t_0) - \eta, T^*(t_0) + \eta),$$

which implies

$$\frac{d}{dt}\mathcal{M}(L^{+}_{\frac{\pi}{3}}(t)) \geq 0 \quad \text{for} \quad t \in (T^*(t_0) - \eta, T^*(t_0) + \eta)$$

by Lemma 6.4 where we use Step B.3 to satisfy the condition $R(t) \geq \underline{R}$. Thanks to the result of Step B.3, we have

$$\begin{aligned}\mathcal{M}(L^{+}_{\frac{\pi}{3}}(t)) &\geq \mathcal{M}(L^{+}_{\frac{\pi}{3}}(T^*(t_0))) \\ &\geq \frac{1}{2}(R(t_0) + 1) - E_1 \quad \text{for} \quad T^*(t_0) \leq t < T^*(t_0) + \eta,\end{aligned}$$

which contradicts to the definition of $T^*(t_0)$. Therefore, we conclude that $T^*(t_0) = \infty$.

E Proof of Corollary 6.2

We next show that the L^2 norm of ϱ in an interval of length $\frac{\pi}{3}$, centered at $-\phi(t)$, decays exponentially after some finite time. For this, we define for each $t \geq 0$ and ω in $[-M, M]$, a functional

$$\Gamma^{-}_{\frac{\pi}{3},\omega}(t) := \int_{L^{-}_{\frac{\pi}{3}}(t)} |\varrho(\theta, \omega, t)|^2 \, d\theta.$$

By Corollary 6.1 and $(\mathcal{H}2)$, we have

$$\inf_{0 \leq t < \infty} R(t) \geq \frac{R_0}{2}.$$

Let $\varepsilon > 0$ be sufficiently small so that

$$\frac{2M}{KR_0} + \frac{4}{K}\frac{M}{R_0^2} + \frac{2\sqrt{2}}{R_0\sqrt{R_0}}\sqrt{\frac{M}{K} + \mu + \varepsilon} - \frac{1}{2} < 0. \tag{E.1}$$

The existence of such ε is guaranteed by the assumption ($\mathcal{H}2$). We set boundary values:

$$\tilde{B}_{-,\omega}(t) := \varrho\Big(\phi(t) + \frac{\pi}{2} + \frac{\pi}{3}, \omega, t\Big), \qquad \tilde{B}_{+,\omega}(t) := \varrho\Big(\phi(t) + \frac{3\pi}{2} - \frac{\pi}{3}, \omega, t\Big).$$

By the same argument as in Section 5.3.2, we have

$$\begin{aligned}\frac{d}{dt}\Gamma^{-}_{\frac{\pi}{3},\omega}(t) &= \dot{\phi}(t)\big(\tilde{B}_{+,\omega}(t)\big)^2 - \dot{\phi}(t)\big(\tilde{B}_{-,\omega}(t)\big)^2 + 2\int_{L^{-}_{\frac{\pi}{3}}(t)} \varrho\partial_t\varrho \, d\theta \\ &=: \dot{\phi}(t)\Big[\big(\tilde{B}_{+,\omega}(t)\big)^2 - \big(\tilde{B}_{-,\omega}(t)\big)^2\Big] + \mathcal{I}_4,\end{aligned} \tag{E.2}$$

where

$$\begin{aligned}\mathcal{I}_4(t) &= -2\int_{L^{-}_{\frac{\pi}{3}}(t)} \varrho\partial_\theta\Big[\varrho\big(\omega - KR\sin(\theta-\phi)\big)\Big] \, d\theta \\ &= -2\int_{L^{-}_{\frac{\pi}{3}}(t)} \Big[(\varrho\partial_\theta\varrho)\big(\omega - KR\sin(\theta-\phi)\big) - KR\varrho^2\cos(\theta-\phi)\Big] \, d\theta \\ &= -\int_{L^{-}_{\frac{\pi}{3}}(t)} (\partial_\theta\varrho^2)\big(\omega - KR\sin(\theta-\phi)\big) \, d\theta + 2KR\int_{L^{-}_{\frac{\pi}{3}}(t)} \varrho^2\cos(\theta-\phi) \, d\theta \\ &=: \mathcal{I}_{41}(t) + \mathcal{I}_{42}(t).\end{aligned} \tag{E.3}$$

Below, we estimate the terms $\mathcal{I}_{4i},\ i = 1, 2$ separately.

- (Estimate of $\mathcal{I}_{41}$): By integration by parts, we have

$$\begin{aligned}\mathcal{I}_{41}(t) = &-\Big[\big(\tilde{B}_{+,\omega}(t)\big)^2\big(\omega - KR\sin(\frac{3\pi}{2} - \frac{\pi}{3})\big) \\ &- \big(\tilde{B}_{-,\omega}(t)\big)^2\big(\omega - KR\sin(\frac{\pi}{2} + \frac{\pi}{3})\big)\Big] \\ &- KR\int_{L^{-}_{\frac{\pi}{3}}(t)} \varrho(\theta,\omega,t)^2\cos(\theta-\phi) \, d\theta \\ &=: \mathcal{I}_{411}(t) + \mathcal{I}_{412}(t).\end{aligned} \tag{E.4}$$

By rearranging the terms in $\mathcal{I}_{411}$, we obtain

$$\mathcal{I}_{411} = -\omega\Big[\big(\tilde{B}_{+,\omega}(t)\big)^2 - \big(\tilde{B}_{-,\omega}(t)\big)^2\Big] - \frac{KR}{2}\Big[\big(\tilde{B}_{+,\omega}(t)\big)^2 + \big(\tilde{B}_{-,\omega}(t)\big)^2\Big]. \tag{E.5}$$

We also combine the terms $\mathcal{I}_{42}$ and $\mathcal{I}_{412}$ and use

$$\cos(\theta - \phi) \le -\sin\frac{\pi}{3} = -\frac{\sqrt{3}}{2} \quad \text{on} \quad L^-_{\frac{\pi}{3}}(t),$$

to obtain

$$\begin{aligned} \mathcal{I}_{42}(t) + \mathcal{I}_{412}(t) &= KR\int_{L^-_{\frac{\pi}{3}}(t)} \varrho(\theta,\omega,t)^2\cos(\theta-\phi)\ d\theta \\ &\le -\frac{KR\sqrt{3}}{2}\Gamma^-_{\frac{\pi}{3},\omega}(t). \end{aligned} \tag{E.6}$$

Finally, in (E.2) we combine (E.3), (E.4), (E.5) and (E.6) to obtain

$$\begin{aligned} \frac{d}{dt}\Gamma^-_{\frac{\pi}{3},\omega}(t) &\le (\dot{\phi}(t) - \omega)\Big[\big(\tilde{B}_{+,\omega}(t)\big)^2 - \big(\tilde{B}_{-,\omega}(t)\big)^2\Big] \\ &\quad - \frac{KR}{2}\Big[\big(\tilde{B}_{+,\omega}(t)\big)^2 + \big(\tilde{B}_{-,\omega}(t)\big)^2\Big] - \frac{KR\sqrt{3}}{2}\Gamma^-_{\frac{\pi}{3},\omega}(t) \\ &\le \Big(-\frac{KR}{2} + |\dot{\phi}(t) - \omega|\Big)\Big[\big(\tilde{B}_{+,\omega}(t)\big)^2 + \big(\tilde{B}_{-,\omega}(t)\big)^2\Big] \\ &\quad - \frac{KR\sqrt{3}}{2}\Gamma^-_{\frac{\pi}{3},\omega}(t). \end{aligned}$$

By Corollary 6.1, there exists $T \ge 0$ such that

$$\dot{R} \le K\mu + K\varepsilon, \quad \text{in } [T,\infty).$$

Using similar arguments as in Lemma 6.2, we have

$$|\dot{\phi}| < \frac{2M}{R_0} + \sqrt{\frac{2K}{R_0}}\sqrt{M + \dot{R}} \le \frac{2M}{R_0} + \sqrt{\frac{2K}{R_0}}\sqrt{M + K(\mu+\varepsilon)},$$

where we used ($\mathcal{H}3$) and Corollary 6.1 to see

$$R(t) > \frac{R_0}{2}, \quad t \in [0, \infty),$$

Thus, by assumption, for any $t \geq T$ we have

$$\begin{aligned}
\frac{d}{dt}\Gamma^-_{\frac{\pi}{3},\omega}(t) &\leq \left(\frac{2M}{R_0} + \sqrt{\frac{2K}{R_0}}\sqrt{M + K\mu + K\varepsilon} + M - K\frac{R_0}{4}\right) \\
&\quad \times \left[\left(\tilde{B}_{+,\omega}(t)\right)^2 + \left(\tilde{B}_{-,\omega}(t)\right)^2\right] - \frac{1}{4}KR_0\Gamma^-_{\frac{\pi}{3},\omega}(t) \\
&= \frac{KR_0}{2}\left(\frac{2M}{KR_0} + \frac{4}{K}\frac{M}{R_0^2} + \frac{2\sqrt{2}}{R_0\sqrt{R_0}}\sqrt{\frac{M}{K} + \mu + \varepsilon} - \frac{1}{2}\right) \\
&\quad \times \left[\left(\tilde{B}_{+,\omega}(t)\right)^2 + \left(\tilde{B}_{-,\omega}(t)\right)^2\right] - \frac{1}{4}KR_0\Gamma^-_{\frac{\pi}{3},\omega}(t) \\
&\leq -\frac{1}{4}KR_0\Gamma^-_{\frac{\pi}{3},\omega}(t),
\end{aligned}$$

where we used (E.1). Then, Gronwall's lemma yields the desired exponential decay:

$$\Gamma^-_{\frac{\pi}{3},\omega}(t) \leq e^{-\frac{KR_0}{4}(t-T)}\Gamma^-_{\frac{\pi}{3},\omega}(T), \quad t \in [T, \infty).$$

On the other hand for $t \in [T, \infty)$ we have

$$\begin{aligned}
&\iint_{L^-_{\frac{\pi}{3}}(t)\times\mathbb{R}} |f|^2 \, d\theta d\omega \\
&= \int_{\mathbb{R}} g^2(\omega) \int_{L^-_{\frac{\pi}{3}}(t)} |\varrho|^2 \, d\theta d\omega = \int_{\mathbb{R}} g^2(\omega)\Gamma^-_{\frac{\pi}{3},\omega}(t) \, d\omega \\
&\leq e^{-\frac{KR(0)}{4}(t-T)} \int_{\mathbb{R}} g^2(\omega)\Gamma^-_{\frac{\pi}{3},\omega}(T) \, d\omega = e^{-\frac{KR(0)}{4}(t-T)} \iint_{L^-_{\frac{\pi}{3}}(T)\times\mathbb{R}} |f(T)|^2 \, d\theta d\omega.
\end{aligned}$$

Thus, we obtain the estimate (6.21). The second estimate in (6.21) is a consequence of the first inequality in (6.21) and Cauchy-Schwarz inequality.

F Dynamics of the Kuramoto-Sakaguchi vector field

In this section, we study analytical properties of integral curves for the Kuramoto-Sakaguchi vector field $\mathcal{X}$ defined by

$$\mathcal{X}(\theta, \omega, t) := \Big(\omega - KR(t)\sin(\theta - \phi(t)), 0, 1\Big). \tag{F.1}$$

Before we study several properties of the integral curves associated with (F.1), we briefly discuss well-posedness of an autonomous ODE. In the sequel, we assume that T_κ is a positive constant satisfying

$$R > \kappa \quad \text{in} \quad [T_\kappa, \infty) \quad \text{and} \quad R > \underline{R} \quad \text{in} \quad [0, \infty), \tag{F.2}$$

for some positive constants $\underline{R}$ and κ.

It follows from Lemma 5.3 and 6.2 that $\mathcal{X}$ is Lipschitz in the given domain. Recall from ($\mathcal{H}$4) just before Proposition 6.1 that

$$\varepsilon_\kappa = \frac{\kappa+1}{\kappa^2}\frac{M}{K} + \frac{(1-\kappa)}{\kappa} < 1. \tag{F.3}$$

Under the assumptions ($\mathcal{H}$1) – ($\mathcal{H}$4), by Proposition 6.2, there exists κ satisfying (F.3) and (F.2) for some T_κ.

We study properties of the integral curves for (F.1) which have been used in the proof of Theorem 3.3.

For a given (θ^*, ω^*) in $\mathbb{T} \times [-M, M]$ and t^* in $[T_\kappa, \infty)$, let $(\theta(t) := (\theta(t; t^*, \theta^*, \omega^*), \omega(t) := \omega(t; t^*, \theta^*, \omega^*))$ be a characteristic curve of (2.15), i.e., it is a solution to the Cauchy problem for the following ODE:

$$\begin{cases} \dot{\theta}(t) = \omega(t) - KR\sin(\theta(t) - \phi(t)), \quad \dot{\omega}(t) = 0, \quad t > T_\kappa, \\ (\theta(t^*), \omega(t^*)) = (\theta^*, \omega^*). \end{cases} \tag{F.4}$$

Since the vector field $\mathcal{X}$ is Lipschitz and $\mathbb{T} \times [-M, M]$ is compact, by Cauchy Lipschitz theorem, characteristics $(\theta(t), \omega(t))$ exists globally and is unique. Below, we will study how the inner product between $e^{\mathrm{i}\theta(t)}$ and $e^{\mathrm{i}\phi(t)}$ can be controlled from above by the solution of an autonomous first order ODE. For this, we first define a *barrier*.

Definition F.1. (Barrier) *For (t^*, p^*) satisfying*

$$t^* \in [T_\kappa, \infty) \quad \textit{and} \quad -\sqrt{1-\varepsilon_\kappa^2} \le p^* \le \sqrt{1-\varepsilon_\kappa^2},$$

the map p:

$$p \ : \ [T_\kappa, t^*] \to (-1, 1),$$

is said to be a barrier through p^* at t^* if it satisfies

$$p(t^*) = p^*, \quad \textit{and} \quad \dot{p} = \kappa K\left(\sqrt{1-p^2} - \varepsilon_\kappa\right)\sqrt{1-p^2}. \tag{F.5}$$

Since the right-hand side of (F.5) is not Lipscthiz at $|p| = 1$, uniqueness is not clear a priori. However, it can be shown that there exists a unique such map $p = p(t)$.

Lemma F.1. *Suppose that supp $g \subset [-M, M]$, and p^*, t^* are positive constants satisfying*

$$-\sqrt{1-\varepsilon_\kappa^2} \le p^* \le \sqrt{1-\varepsilon_\kappa^2}, \qquad t^* \ge T_\kappa. \tag{F.6}$$

Then, the barrier $p = p(t)$ through p^* at t^* is unique and satisfies

$$-\sqrt{1-\varepsilon_\kappa^2} \le p(t) \le \sqrt{1-\varepsilon_\kappa^2}, \quad \textit{for all } t.$$

Proof. Let $\varepsilon > 0$ be a positive constant satisfying

$$\sqrt{1-\varepsilon_\kappa^2} + \varepsilon < 1.$$

Recall F_κ defined in (**??**):

$$F_\kappa(q) := \kappa K(\sqrt{1-q^2} - \varepsilon_\kappa)\sqrt{1-q^2}, \quad q \in [-1,1].$$

Choose a Lipschitz function $\tilde{F}_\kappa$ compactly supported on $(-1,1)$ that coincides with F_κ in $[-\sqrt{1-\varepsilon_\kappa^2}, \sqrt{1-\varepsilon_\kappa^2}]$. Since $\tilde{F}_\kappa$ has compact support, it follows from Cauchy-Lipschitz theorem that the equation

$$\dot{q} = \tilde{F}_\kappa(q),$$

has a unique solution. Moreover, since

$$q_1(t) = -\sqrt{1-\varepsilon_\kappa^2} \quad \text{and} \quad q_2(t) = \sqrt{1-\varepsilon_\kappa^2},$$

are solutions, uniqueness of ODE implies that any solution satisfying

$$-\sqrt{1-\varepsilon_\kappa^2} \le q(t^*) \le \sqrt{1-\varepsilon_\kappa^2},$$

satisfies

$$-\sqrt{1-\varepsilon_\kappa^2} \le q(t) \le \sqrt{1-\varepsilon_\kappa^2}, \quad \text{for all } t.$$

This complete the proof. □

Next we study a quantitative growth estimate to be used in Corollary F.2.

Lemma F.2. *Let* p^*, ε, *and* t^* *be positive constants satisfying*

$$0 < \varepsilon < \sqrt{1-\varepsilon_\kappa^2}, \qquad p^* = \sqrt{1-\varepsilon_\kappa^2} - \varepsilon, \qquad D(\varepsilon,\kappa) < t^* - T_\kappa. \tag{F.7}$$

Then, there exists a unique constant $d < D(\varepsilon,\kappa)$ *such that the barrier* p *through* p^* *at* t^* *satisfies*

$$p(t^*-d) = -\sqrt{1-\varepsilon_\kappa^2} + \varepsilon,$$

Here, $D(\varepsilon,\kappa) = \frac{2(\sqrt{1-\varepsilon_\kappa^2}-\varepsilon)}{F_\kappa(\sqrt{1-\varepsilon_\kappa^2}-\varepsilon)}$ *as given in* (**??**)

Proof. It follows from Lemma F.1 that we have

$$-\sqrt{1-{\varepsilon_\kappa}^2} < p(t) < \sqrt{1-{\varepsilon_\kappa}^2}, \quad \text{for all } t.$$

Since

$$\dot{p} = F_\kappa(p) \geq F_\kappa\big(\sqrt{1-\varepsilon_\kappa^2} - \varepsilon\big) > 0,$$

whenever p is contained in the interval

$$\Big[-\sqrt{1-\varepsilon_\kappa^2} + \varepsilon, \sqrt{1-\varepsilon_\kappa^2} - \varepsilon \Big],$$

there exist a unique d such that

$$p(t^*-d) = -\sqrt{1-\varepsilon_\kappa^2} + \varepsilon. \tag{F.8}$$

Then, we have

$$p(t^*) - p(t^*-d) = \int_{t^*-d}^{t*} \dot{p}\; dt \geq \int_{t^*-d}^{t*} F_\kappa\big(\sqrt{1-\varepsilon_\kappa^2} - \varepsilon\big) dt = F_\kappa\big(\sqrt{1-\varepsilon_\kappa^2} - \varepsilon\big) d.$$

This yields

$$d < \frac{p(t^*) - p(t^* - d)}{F_\kappa\big(\sqrt{1-\varepsilon_\kappa^2} - \varepsilon\big)},$$

and the desired estimate follows from (F.7) and (F.8). $\square$

Lemma F.3. *Let t^*, θ^*, ω^* and p^* be constants satisfying*

$$p* \in \big[-\sqrt{1-\varepsilon_\kappa^2}, \sqrt{1-\varepsilon_\kappa^2}\big], \quad \omega^* \in [-M, M], \quad t^* \geq T_\kappa$$
$$\text{and} \quad \cos(\theta^* - \phi(t^*)) \leq p^*.$$

Then, the characteristics $(\theta(t), \omega(t))$ through (θ^*, ω^*) at t^* satisfies

$$\cos(\theta(t) - \phi(t)) \leq p(t) \quad \forall t \in [T_\kappa, t^*],$$

where p is the barrier through p^* at t^*.

Proof. Let $P : [T_\kappa, t^*] \to [-1, 1]$ be defined by the following relation:

$$P(t) := \cos(\theta(t) - \phi(t)).$$

It follows from (F.4) and Lemma 5.2 that for $[T_\kappa, t^*]$,

$$\begin{aligned}
\dot{P}(t) &= -\big(\dot{\theta}(t) - \dot{\phi}(t)\big)\sin(\theta(t) - \phi(t)) \\
&= -\big(\omega(t) - KR(t)\sin(\theta(t) - \phi(t)) - \dot{\phi}(t)\big)\sin(\theta(t) - \phi(t)) \\
&= KR(t)\sin^2(\theta(t) - \phi(t)) + [\dot{\phi}(t) - \omega(t)]\sin(\theta(t) - \phi(t)) \\
&= KR(t) - KR(t)\cos^2(\theta(t) - \phi(t)) + [\dot{\phi}(t) - \omega(t)]\sin(\theta(t) - \phi(t)) \\
&\geq KR(t)\big(1 - \cos^2(\theta(t) - \phi(t))\big) - (|\dot{\phi}(t)| + M)\sqrt{1 - \cos^2(\theta(t) - \phi(t))} \\
&> K\kappa\big(1 - |P(t)|^2\big) - \left[\left(1 + \frac{1}{\kappa}\right)M + K(1-\kappa)\right]\sqrt{1 - |P(t)|^2},
\end{aligned}$$

where we used (5.17) and (F.2) in the last line. Thus, we have

$$\dot{P} > K\kappa(\sqrt{(1-P^2)} - \varepsilon_\kappa)\sqrt{1-P^2}. \tag{F.9}$$

Thus, since $P(t^*) \leq p^*$, the desired result follows from (F.5), (F.9) and Lemma F.1. $\square$

As a direct application of Lemma F.3, we have the following two corollaries.

Corollary F.1. *Let δ, θ^*, ω^*, and t^* be positive constants such that*

$$0 < \delta < \sqrt{1-\varepsilon_\kappa^2}, \quad \omega^* \in [-M, M], \quad t^* \geq T_\kappa \quad \textit{and} \quad \cos(\theta^* - \phi(t^*)) \leq -\delta.$$

Then, the characteristics $(\theta(t), \omega(t))$ through (θ^*, ω^*) at t^* satisfies

$$\cos(\theta(t) - \phi(t)) \leq -\delta, \quad \forall t \in [T_\kappa, t^*].$$

Proof. Let p be the barrier through $-\delta$ at t^*. By Lemma F.1 and (F.5), we have that p is nondecreasing, and since

$$\cos(\theta(t^*) - \phi(t^*)) \leq p(t^*) = -\delta,$$

the desired result follows from Lemma F.3 using $-\delta = -p^*$. $\square$

Corollary F.2. *Let ε, θ^*, ω^*, and t^* be positive constants satisfying*

$$\begin{aligned} &0 < \varepsilon < \sqrt{1-\varepsilon_\kappa^2}, \quad D(\varepsilon, \kappa) < t^* - T_\kappa, \quad \omega^* \in [-M, M], \\ &\textit{and} \quad \cos(\theta^* - \phi(t^*)) \leq \sqrt{1-\varepsilon_\kappa^2} - \varepsilon. \end{aligned}$$

Then, there exists a positive constant d satisfying

$$d < D(\varepsilon, \kappa) \quad \textit{and} \quad \cos(\theta(t^* - d) - \phi(t^* - d)) \leq -\sqrt{1 - \varepsilon_\kappa^2} + \varepsilon.$$

where $D(\varepsilon, \kappa)$ is a positive constant defined in (**??**)*, and $(\theta(t), \omega(t))$ is the characteristics passing through (θ^*, ω^*) at t^*.*

Proof. Let p be the barrier through $\sqrt{1-\varepsilon_\kappa^2} - \varepsilon$ at t^*. By Lemma F.3, we have

$$\cos(\theta(t) - \phi(t)) \leq p(t), \qquad t \in [T_\kappa, t^*],$$

and the desired result follows by (F.8). □

G Weak Kam structure for kinetic equations

The objective of this appendix is to introduce a heuristic framework in which several kinetic equation have a formal weak KAM structure. The finite dimensional version of this structure is described in section 1.4 of the introduction. We begin by adapting the action functionals in the space of paths of measures, to phase space. In d dimensions a point in phase space is given by a couple $(x, v) \in \mathbb{R}^d \times \mathbb{R}^d$. A delta measure centered at a given point (x, v) represents a particle having position x and velocity v. Hence, if the map $t \to (x(t), v(t))$ represents the motion of a real particle, then we must have $\dot{x}(t) = v(t)$. When we consider a curve of measures $t \to f_t$ having velocity field $t \to v_t = (v_{t,x}, v_{t,v})$, i.e.

$$\partial_t f = -\mathrm{div}_x(f v_{t,x}) - \mathrm{div}_v(f v_{t,v}),$$

and minimizing our action functionals, we would like that any integral curve $t \to (x(t), v(t))$, corresponding to the flow induced by v_t, satisfies the condition $\dot{x} = v(t)$. That is, we want the curve of measures $t \to f_t$ to represent the motion of real particles in phase space. We will achieve this by using a penalization argument . We describe such argument in the general setting of the tangent space to a Riemannian manifold (M, g). This will allow us to consider a bigger family of equations and include the Cucker-Smale model, the stochastic Vicsek model, and the Kuramoto-Sakaguchi equation (see sections 1.4 and 5 of the introduction, there we explain how such heuristic framework helped us achieve or concentration result in Theorem 3.3 of chapter 5).

G.1 Penalized action functionals

Let (M, g) be a Riemannian manifold and let W be a vector field in M. For each $x \in M$ we let $\mathcal{F}(x)$ be a closed, convex set in T_xM. Additionally, let Λ be a positive number and μ and ν be measures in $\mathbb{P}(M)$. We consider the problem of minimizing.

$$\int_0^\tau \int \left[\frac{1}{2}|v_t|^2 + W \cdot v_t + \frac{\Lambda}{2} \inf_{w \in \mathcal{F}(x)} \mid v_t - w \mid^2 \right] \rho_t \, dxdt, \tag{G.10}$$

among all measured valued maps $t \to \rho_t$ and vector fields $t \to v_t$ from $[0, \tau]$ to $\mathbb{P}_2(M)$ such that $\rho_0 \, dx = \mu$, $\rho_\tau \, dx = \nu$, and

$$\frac{d}{dt} \int \zeta\rho \, dx = \int \nabla\zeta v_t \rho_t \, dx \quad \forall \zeta \in C^\infty(M) \quad \text{and} \quad \forall t \in [0, \tau]. \tag{G.11}$$

Here, all the norms, inner products, gradients and divergences will be taken with respect to g. Additionally, We assume the assignation $x \to \mathcal{F}(x)$, is well behaved. By this we mean that the map $P : TM \to TM$, given by

$$P(x,v) = \left(x, \mathrm{argmin}_{w \in \mathcal{F}(x)} \frac{1}{2} \mid v - w \mid^2 \right),$$

is differentiable.

G.2 Non-coercive Hamiltonians

To understand minimizers of (G.10) when $\Lambda \to \infty$, it will be convenient to compute the Hamiltonian associated with the Lagrangian

$$L(x,v) = \frac{1}{2}|v|^2 + W \cdot v + \frac{\Lambda}{2} \inf_{w \in \mathcal{F}(x)} \mid v - w \mid^2 .$$

For this purpose we note that if H is the Hamiltonian associated with L, then

$$\begin{aligned} H(x,p) &= \sup_v pv - \frac{1}{2}v^2 - W \cdot v - \frac{\Lambda}{2} \inf_{w \in \mathcal{F}(x)} \left(v - w\right)^2 \\ &= \sup_v \sup_{w \in \mathcal{F}(x)} pv - \frac{1}{2}v^2 - W \cdot v - \frac{\Lambda}{2} \inf \mid v - w \mid^2 \\ &= \sup_{w \in \mathcal{F}(x)} \sup_v pv - \frac{1}{2}v^2 - W \cdot v - \frac{\Lambda}{2} \mid v - w \mid^2 \\ &= \sup_{w \in \mathcal{F}(x)} p\frac{p - W + \Lambda w}{1+\Lambda} - \frac{1}{2}\left|\frac{p - W + \Lambda w}{1+\Lambda}\right|^2 - W \cdot \frac{p - W + \Lambda w}{1+\Lambda} \\ &\qquad - \left|\frac{p - W + \Lambda w}{1+\Lambda} - w\right|^2 . \end{aligned} \tag{G.12}$$

Here, we have used the relationships

$$p = v_p + W + \Lambda(v_p - w),$$

and

$$v_p = \frac{p - W + \Lambda w}{1+\Lambda}, \tag{G.13}$$

where for each p, v_p is the element in T_xM maximizing

$$v \to pv - \frac{1}{2}v^2 - W \cdot v - \frac{\Lambda}{2} \inf_{w \in \mathcal{F}(x)} \big(v - w\big)^2.$$

Now, when we let $\Lambda \to \infty$ in (G.12), we obtain, $H \to \mathcal{H}$, where

$$\begin{aligned}
\mathcal{H}(x,p) &= \sup_{w \in \mathcal{F}(x)} pw - \frac{1}{2}|w|^2 - w \cdot W \\
&= \frac{1}{2}|p - W|^2 + \sup_{w \in \mathcal{F}(x)} -\frac{\mid p - W - w \mid^2}{2} \\
&= \frac{1}{2}|p - W|^2 - \inf_{w \in \mathcal{F}(x)} \frac{\mid p - W - w \mid^2}{2} \\
&= \frac{1}{2}|p - W|^2 - \frac{\mid p - W - P(x, p - W) \mid^2}{2} \\
&= pP(x, p - W) - P(x, p - W) \cdot W - \frac{1}{2} \mid P(x, p - W) \mid^2 .
\end{aligned}$$

Then,

$$\begin{aligned}
\mathcal{L}(x,v) &= \sup_p pv - pP(x, p - W) - P(x, p - W) \cdot W + \frac{1}{2} \mid P(x, p - W) \mid^2 \\
&= \begin{cases} \frac{1}{2} \mid P(x, p - W) \mid^2 - P(x, p - W) \cdot W & \text{if} \quad P(x, p - W) = v, \\ \infty & \text{otherwise.} \end{cases}
\end{aligned}$$

Consequently, if $\mathcal{L}(x,v)$ is finite and v and p are conjugates via the Legendre transform (see section 1.4 and relationship (G.13)), then we must have

$$v = P(x, p - W).$$

Here, we are abusing notation to regard the covector p as a vector. To do this we are using the canonical isomorphism between TM^* and TM induced by g. From now on we will adopt this convention.

G.3 Optimal velocity fields as an infinite dimensional Legendre transform

Suppose the curve $t \to \rho_t$ with velocity field $t \to v_t$, is a minimizer of (G.10) and fix $t \geq 0$. As in section 2.2 of the introduction, we note that if $t \to v_t, \rho_t$ satisfies constraint (G.11), so does $v_t + sA$, provided $\operatorname{div}(A\rho_t) = 0$. Consequently, by minimality we must have

$$0 = \frac{d}{ds}\bigg|_{s=0} \int \Big[L(x, v_t + sA)\Big]\rho_t \, dx = \int \nabla_v L(x, v_t) A \, dx.$$

Since this hold for any A with the property that $\operatorname{div}(A\rho_t) = 0$, by the Helmholtz-Hodge Theorem there exists $\nabla\varphi_t$ with the property that

$$\nabla_v L(x, v_t) = \nabla\varphi_t. \tag{G.14}$$

Using Legendre duality we conclude

$$\nabla_p H(x, \nabla\varphi_t) = v_t.$$

As $\Lambda \to \infty$, the discussion of the previous section yields

$$v_t = P(x, \nabla\varphi_t - W).$$

In the case where $\mathcal{F}(x)$ is a subspace, this is just the orthogonal projection of $\nabla\varphi_t - W$ into $\mathcal{F}(x)$.

Relationship (G.14) can be regarded as an infinite dimensional Legendre transform. Indeed, in finite dimensions, if v_p is the Legendre conjugate of p then,

$$v_p = \operatorname{argmax}_v \, pv - L(x, v).$$

On the other hand, when we use the formal Riemannian structure of Felix Otto, with

$$T_\rho \mathbb{P}_2(M) = \overline{\left\{ \nabla\varphi \,:\, \int |\nabla\varphi|^2 \rho \, dx < \infty \right\}}^{L^2(\rho)}$$

and we consider the Legendre conjugate of a covector $\nabla\varphi$ we maximize

$$\omega \to \int \nabla\varphi \nabla\omega \rho \, dx - \int L(x, \nabla\omega)\rho \, dx.$$

If ω is a maximizer of the above map, then we must have that for any function α,

$$\begin{aligned} 0 &= \frac{d}{ds} \int \nabla\varphi \nabla(\omega + s\alpha)\rho \, dx - \int L(x, \nabla(\omega + s\alpha))\rho \, dx \\ &= \int \nabla\varphi \nabla\alpha\rho \, dx - \int \nabla_v L(x, \nabla\omega)\nabla\alpha\rho \, dx. \end{aligned}$$

Since α was arbitrary, we must have

$$\nabla\varphi = \nabla_v L(x, \nabla\omega),$$

which is equivalent to (G.14).

G.4 Extremals of the action

In this section, we show a heuristic argument that characterizes minimizers of (G.10) and their limiting properties as $\Lambda \to \infty$. The main idea is to generalize the techniques of section 3 of the introduction to compute the first variation of paths in $\mathbb{P}_2(M)$ minimizing (G.10). Let $t \to \rho_t$, be a minimizer of (G.10), and let $t \to \varphi_t$ be the potential generating the corresponding optimal velocity field indexed in $[0, \tau]$ as described in the previous section so that

$$\frac{d}{dt} \int \zeta \rho_t \, dx = \int \nabla\zeta \nabla_p H(x, \nabla\varphi_t)\rho_t \, dx \quad \forall \zeta \in C^\infty(M) \quad \text{and} \quad \forall t \in [0, \tau].$$

(see section 5.2.3).

We will we show that $t \to \varphi_t$, satisfies

$$\partial_t \varphi_t + H(x, \nabla \varphi_t) = 0 \quad \text{in} \quad [0, \tau],$$

Here, H is the Hamiltonian corresponding to the Lagrangian

$$L(x, v) = \frac{1}{2}|v|^2 + v \cdot W + \frac{\Lambda}{2} \inf_{w \in \mathcal{F}(x)} \mid v - w \mid^2 .$$

Hence, by the discussion Section (G.2), we get that when $\Lambda \to \infty$, minimizers satisfy

$$\partial_t \varphi + \nabla \varphi P(x, \nabla \varphi - W) - W \cdot P(x, \nabla \varphi - W) - \frac{1}{2} \mid P(x, \nabla \varphi - W) \mid^2 = 0,$$

and

$$\frac{d}{dt} \int \zeta \rho_t \, dx = \int \nabla \zeta P(x, \nabla \varphi - W) \rho_t \, dx \quad \forall \zeta \in C^\infty(M) \quad \text{and a.e } t \quad \text{in} \quad [0, \tau].$$

In order to see this, we perturb such minimizers by building a vector field along the minimizing path, (as we did in section 3 of the introduction). That is, for each t in $[0, \tau]$ we consider a potential ω_t and the vector field $\nabla \omega_t$ induced by it. We require ω_t to be identically 0 in the complement of a compact subset of $(0, \tau)$, so that the variation fixes the end points of the optimal path.

Indeed, for each s, we let $t \to \rho_{t,s}$ and $t \to \nabla_p H(x, \nabla \varphi_{t,s})$ satisfy

$$\frac{d}{dt} \int \zeta \rho_{t,s} dx = \int \nabla \zeta \nabla_p H(x, \nabla \varphi_{t,s}) \rho_{t,s} dx \quad \forall \zeta \in C^\infty(M),$$

and

$$\frac{d}{ds}\bigg|_{s=0} \int \zeta \rho_{t,s} dx = \int \nabla \zeta \nabla \omega_t \rho_{t,s} dx \quad \forall \zeta \in C^\infty(M).$$

Then, by minimality of the path $t \to \rho_t$, when $s = 0$, we must have

$$\begin{aligned}
0 = \frac{d}{ds} \int_0^\tau \int & L(x, \nabla_p H(x, \nabla \varphi_{t,s}) \rho_{t,s}\, dx \\
&= \int_0^\tau \bigg[\int \nabla_v L(x, \nabla_p H(x, \nabla \varphi_{t,s})) \partial_s \nabla_p H(x, \nabla \varphi_{t,s}) \rho_{t,s}\, dx \\
&+ \int L(x, \nabla_p H(x, \nabla \varphi_{t,s})) \partial_s \rho_{t,s}\, dx \bigg]\, dt \\
&= \int_0^\tau \bigg[\int \nabla \varphi_{t,s} \partial_s \nabla_p H(x, \nabla \varphi_{t,s}) \rho_{t,s}\, dx \\
&+ \int L(x, \nabla_p H(x, \nabla \varphi_{t,s})) \partial_s \rho_{t,s}\, dx \bigg]\, dt \\
&= -\int_0^\tau \bigg[\int \nabla \partial_s \varphi_{t,s} \nabla_p H(x, \nabla \varphi_{t,s}) \rho_{t,s}\, dx \\
&+ \frac{d}{ds} \int \nabla \varphi_{t,s} \nabla_p H(x, \nabla \varphi_{t,s}) \rho_{t,s}\, dx \\
&- \int \nabla \varphi_{t,s} \nabla_p H(x, \nabla \varphi_{t,s}) \partial_s \rho_{t,s}\, dx \\
&+ \int L(x, \nabla_p H(x, \nabla \varphi_{t,s})) \partial_s \rho_{t,s}\, dx \bigg]\, dt \\
&= \int_0^\tau \bigg[\int -\partial_s \varphi_{t,s} \partial_t \rho_{t,s}\, dx \\
&+ \frac{d}{ds} \int \varphi_{t,s} \partial_t \rho_{t,s}\, dx - \int H(x, \nabla \varphi_{t,s}) \partial_s \rho_{t,s}\, dx \bigg]\, dt \\
&= \int_0^\tau \bigg[\varphi_{t,s} \partial_{ts} \rho_{t,s}\, dx - \int H(x, \nabla \varphi_{t,s}) \partial_s \rho_{t,s}\, dx \bigg]\, dt \\
&= -\int_0^\tau \Big(\partial_t \varphi_{t,s} + H(x, \nabla \varphi_{t,s}) \Big) \partial_s \rho_{t,s}\, dx\, dt \\
&= -\int_0^\tau \int \bigg[\nabla \omega_t \nabla [\partial_t \varphi_t + H(x, \nabla \varphi_t)] \rho_t\, dx \bigg]\, dt.
\end{aligned}$$

Since $t \to \omega_t$ was arbitrary, the desired result follows.

G.5 Minimizing movement scheme of the entropy

In this section, we consider an infinite dimensional variant of the method described in section 1.5 of the introduction to build calibrated curves of a KAM function starting at a point x_0. By doing this, in the next section, we will be able to regard a large family of kinetic equations as calibrated curves of KAM functions.

Let ρ_0 be a probability density in $\mathbb{P}_2(M)$. We provide a heuristic argument to characterize minimizers of

$$\{\rho_t\}_{t\in[0,\tau]} \to \int_0^\tau \frac{1}{2}\int \left[|v_t|^2 + W\cdot v_t + \frac{\Lambda}{2}\inf_{\omega\in\mathcal{F}(x)}|v_t-\omega|^2\right]\rho_t\, dxdt, \qquad \text{(G.15)}$$

as well as their properties when $\Lambda\to\infty$. Here, the vector fields $t\to v_t$ satisfy constraint (G.11) and ρ_τ is a free parameter in the minimization.

In section G.3, we saw that any optimal velocity field has, for each t in $[0,\tau]$, a function φ_t such that

$$v_t = \nabla_p H(x,\nabla\varphi_t).$$

In section G.4, we found that any potential $t\to\varphi_t$ corresponding to an optimal path satisfies

$$\partial_t\varphi + H(x,\nabla\varphi_t) = 0,$$

and

$$\frac{d}{dt}\int\zeta\rho_t dx = \int\nabla\zeta\nabla_p H(x,\nabla\varphi_t)\rho_t\, dx \quad \forall\zeta\in C^\infty(M) \quad \text{and} \quad \forall t\in[0,\tau].$$

In this section, we will show that minimizers of (G.15) satisfy

$$\varphi_\tau = -\log \rho_\tau - 1 + c,$$

for some constant c.

We proceed by perturbing the path $t \to \rho_t$, in the same way as the previous section. Let $t \to \nabla_p H(x, \nabla\varphi_t)$ be the optimal velocity field associated to the optimal paths (see section G.3). For each t we choose a function ω_t. We require these function to be identically 0 in the complement of a compact subset of $(0, \tau]$. This generates for each s a path $t \to \rho_{t,s}$, as in section G.4, with the difference that the end point $\rho_{\tau,s}$ is free.

$$\begin{aligned}
0 = \frac{d}{ds}\int \rho_{\tau,s}\log\rho_{\tau,s}\,dx + \int_0^\tau \int L(x, \nabla_p H(\nabla\varphi_{t,s})\rho_{t,s}\,dxdt \\
&= \int (\log\rho_{\tau,s} + 1)\partial_s\rho_{h,s}\,dx \\
&+ \int \varphi\partial_{ts}\rho_{t,s}dx - \int_0^1 H(x, \nabla\varphi_{t,s})\partial_s\rho_{t,s}\,dx \\
&= \int (\log\rho_{\tau,s} + 1 + \varphi_\tau)\partial_s\rho_{\tau,s}\,dx \\
&= \int \nabla(\log\rho_{\tau,s} + 1 + \varphi_\tau)\nabla\omega_\tau\rho_{\tau,s}\,dx.
\end{aligned}$$

(Here, in the second equality we have integrated by parts in t and used the fact that by construction $\partial_s\varphi_{0,s} = \partial_s\rho_{0,s} = 0$.)

Since ω_τ was arbitrary, the desired result follows.

G.6 Examples

The Cucker-Smale model

We let $(M,g) = \mathbb{R}^d \times \mathbb{R}^d$, with the standard product metric. Additionally, we set $W = 0$, $\mathcal{F}(x,v) = \{v\} \times \mathbb{R}^d \subset T_{(x,v)}(\mathbb{R}^d \times \mathbb{R}^d)$. Then, by section G.3, when $\Lambda \to \infty$, we have

$$P\big((x,v), \nabla\varphi\big) = (v, \nabla_v \varphi).$$

Hence, by section G.4, optimal paths satisfy

$$\begin{aligned}\frac{d}{dt}\int \zeta f_t d\omega d\sigma &= \int \nabla\zeta P\big((x,v), \nabla\varphi_t\big) f_t \, d\omega d\sigma \\ &= \int \big[v\nabla_x\zeta + \nabla_v\zeta\nabla_v\varphi_t\big] f_t \, d\omega d\sigma,\end{aligned}$$

for any $\zeta \in C^\infty(M)$ and

$$\begin{aligned}0 &= \partial_t\varphi + \nabla\varphi P(x,\nabla\varphi) - \frac{1}{2} \mid P(x,\nabla\varphi) \mid^2 \\ &= \partial_t\varphi + v\cdot\nabla_x\varphi + \frac{1}{2}|\nabla_v\varphi|^2 - \frac{1}{2}|v|^2.\end{aligned}$$

Then, proceeding as in section G.5 when one minimizes

$$\begin{aligned}\{f_t\}_{t\in[0,\tau)} \to \frac{1}{4}\int & \psi(|x-x^*|)|v-v^*|^2 f_\tau(x,v) f_\tau(x^*,v^*) dxdvdx^*dv^* \\ &+ \int_0^\tau \left(\int L(x,v_t) f_t \, d\omega d\sigma\right) dt,\end{aligned}$$

with f_0 fixed, one finds the optimality conditions

$$\begin{aligned}0 &= \int (\varphi_\tau - \frac{1}{4}\int \psi(|x-x^*|)|v-v^*|^2 dx^*dv^*)\partial_s f_{\tau,s} \, dxdv \\ &= \int \nabla(\varphi_\tau + \int \psi(|x-x^*|)|v-v^*|^2 dx^*dv^*)\nabla\omega_\tau f_{\tau,s} \, dxdv.\end{aligned}$$

Thus,

$$\varphi_\tau(x,v) = -\frac{1}{2}\int \psi(|x-x^*|)|v-v^*|^2 f_\tau(x^*,v^*)\, dx^*dv^* + c,$$

for some constant c.

Consequently, when one follows the minimal movement scheme and lets $\Lambda \to \infty$, one should obtain solutions to

$$\begin{aligned}
\partial_t f &= -\mathrm{div}\Big(fP\big(x, \nabla \frac{1}{2}(\int \psi(|x-x^*|)|v-v^*|^2 dx^*dv^*)\big)\Big) \\
&= -\mathrm{div}_x\big(vf\big) - K\mathrm{div}_v\big(f\nabla_v \frac{1}{2}\int \psi(|x-x^*|)|v-v^*|^2 dx^*dv^*\big) \\
&= -\mathrm{div}_x\big(vf\big) - K\mathrm{div}_v\big(fF_a(f)\big),
\end{aligned}$$

where

$$F_a(f) = \int \psi(|x-x^*|)(v^*-v)f\; dx^*dv^*.$$

Stochastic Vicsek model

We let $(M,g) = \mathbb{R}^d \times \mathbb{S}^{d-1}$ with the standard product metric. Additionally, we set $W = 0$ and $\mathcal{F}(x,\omega) = \{\omega\} \times \{T_\omega \mathbb{S}^{d-1}\} \subset T_{(x,\omega)}(\mathbb{R}^d \times \mathbb{S}^{d-1})$.

Then, by section G.3, when $\Lambda \to \infty$ we have

$$P(x, \nabla\varphi) = (\omega, \nabla_\omega \varphi).$$

Here, ∇_ω denotes the gradient in $\mathbb{S}^{d-1}$ with the standard metric.

Hence, by section G.4, optimal paths satisfy

$$\frac{d}{dt}\int \zeta f_t \, dxd\omega = \int \nabla\zeta P(x,\nabla\varphi_t) f_t \, dxd\omega = \int \nabla_x \zeta \omega f_t \, dxdv$$
$$+ \int \nabla_\omega \zeta \nabla_\omega \varphi_t f_t \, dxdv$$
$$\forall \zeta \in C^\infty(M),$$

and

$$0 = \partial_t\varphi + \nabla\varphi P(x,\nabla\varphi) - \frac{1}{2} \mid P(x,\nabla\varphi) \mid^2 = \partial_t\varphi + \omega\cdot\nabla_x\varphi + \frac{1}{2} \mid \nabla_\omega\varphi \mid^2 - \frac{1}{2}.$$

Then, proceeding as in section G.5 when one minimizes

$$\{f_t\}_{t\in[0,\tau)} \to \int \left(f_\tau \log f_\tau - \mid J_{f_\tau}(x) \mid \right) dxdv + \int_0^\tau \left(\int L(x,v_t) f_t dv d\omega \right) dt,$$

with f_0 fixed, one finds the optimality condition

$$\begin{aligned} 0 &= \int (\varphi_\tau + \log f_\tau + 1 - \omega\cdot\Omega_{f_\tau}(x))\partial_s f_{\tau,s} dv d\omega \\ &= \int \nabla\omega_\tau \nabla(\varphi_\tau + \log f_\tau + 1 - \omega\cdot\Omega_{f_\tau}(x)) dv d\omega. \end{aligned}$$

Here

$$J_f(x) = \int \omega f(x,\omega) d\omega,$$

and

$$\Omega_f(x) = \frac{J_f(x)}{\mid J_f(x) \mid}.$$

Thus,

$$\varphi_\tau = -\log \frac{f_\tau}{e^{\omega\cdot\Omega_{f_\tau}}} + c,$$

for some constant c.

Consequently, when one follows the minimal movement scheme and lets $\Lambda \to \infty$, one should obtain solutions to

$$
\begin{aligned}
\partial_t f &= -\mathrm{div}(fP(x, -\nabla \log \frac{f_\tau}{e^{\omega\cdot\Omega_{f_\tau}}})) \\
&= -\mathrm{div}_x(\omega f) + \mathrm{div}_\omega\big(f\nabla \log \frac{f}{e^{\omega\cdot\Omega_f}}\big) \\
&= -\mathrm{div}_x(\omega f) + \Delta_\omega f - \mathrm{div}_\omega(f\nabla_\omega \omega \cdot \Omega_f).
\end{aligned}
$$

Kuramoto Synchronization model

We let $(M, g) = \mathbb{R}\times\mathbb{S}^1$, with the standard product metric. Additionally, we set $W(\omega, \sigma) = (0, -\omega\sigma^\perp) \in T_{(\omega,\sigma)}(\mathbb{R} \times \mathbb{S}^1)$, $\mathcal{F}(\omega, \sigma) = \{0\} \times \{T_\sigma\mathbb{S}^1\} \subset T_{(\omega,\sigma)}(\mathbb{R} \times \mathbb{S}^1)$. Then, by section G.3, when $\Lambda \to \infty$, we have

$$P(x, \nabla\varphi) = (0, \nabla_\sigma\varphi).$$

Here, ∇_σ denotes the gradient in $\mathbb{S}^1$ with the standard metric.

Hence, by section G.4, optimal paths satisfy

$$
\begin{aligned}
\frac{d}{dt}\int \zeta f_t d\omega d\sigma &= \int \nabla\zeta P\big(x, \nabla\varphi_t + (0, \omega\sigma^\perp)\big) f_t \, d\omega d\sigma \\
&= \int \nabla_\sigma\zeta \cdot [\omega\sigma^\perp + \nabla_\sigma\varphi\big] f_t \, d\omega d\sigma,
\end{aligned}
$$

and

$$
\begin{aligned}
0 &= \partial_t\varphi + \nabla\varphi P(x, \nabla\varphi + (0, \omega\sigma^\perp)) + (0, \omega\sigma^\perp) \cdot P(x, \nabla\varphi + (0, \omega\sigma^\perp)) \\
&\quad - \frac{1}{2} \mid P(x, \nabla\varphi + (0, \omega\sigma^\perp)) \mid^2 \\
&= \partial_t\varphi + \frac{1}{2}|\nabla_\sigma\varphi|^2 + \omega\sigma^\perp \cdot \nabla_\sigma\varphi + \frac{1}{2}\omega^2,
\end{aligned}
$$

for any $\zeta \in C^\infty(M)$ and t in $[0, \infty)$.

Then, proceeding as in section G.5 when one minimizes

$$\{f_t\}_{t\in[0,\tau)} \to \frac{K}{2}\left(1 - |J_{f_\tau}|^2\right) + \int_0^\tau \left(\int L(x, v_t) f_t \, d\omega d\sigma \right) dt,$$

with f_0 fixed, one finds the optimality condition

$$\begin{aligned} 0 &= \int (\varphi_\tau - K\sigma \cdot J_f) \partial_s f_{\tau,s} \, d\omega d\sigma \\ &= \int \nabla(\varphi_\tau - K\sigma \cdot J_{f_\tau}) \nabla \omega_\tau f_\tau \, d\omega d\sigma. \end{aligned}$$

Here,

$$J_f(x) = \int \sigma f(\sigma, \omega) \, d\sigma d\omega,$$

Thus,

$$\varphi_\tau = K\sigma \cdot J_{f_\tau} + c,$$

for some constant c.

Consequently, when one follows the minimal movement scheme and lets $\Lambda \to \infty$, one should obtain solutions to

$$\begin{aligned} \partial_t f &= -\mathrm{div}(f P(x, K\nabla\sigma \cdot J_{f_\tau} + (0, \omega\sigma^\perp))) \\ &= -K\mathrm{div}_\sigma\big(f \nabla_\sigma \sigma \cdot J_{f_\tau}\big) - K\mathrm{div}_\sigma\big(f\omega\sigma^\perp\big), \end{aligned}$$

which is the Kuramoto Sakaguchi equation in $\mathbb{R} \times \mathbb{S}^1$.

Bibliography

[1] Acebron, J.A., Bonilla, L. L., Pérez Vicente, C.J.P., Ritort, F. and Spigler, R.: *The Kuramoto model: A simple paradigm for synchronization phenomena.* Rev. Mod. Phys. **77**, 137-185 (2005).

[2] Amadori, D., Ha, S.-Y. and Park, J.: *On the global well-posedness of BV weak solutions for the Kuramoto-Sakaguchi equation.* Submitted.

[3] Aldana M., and Huepe C., *Phase transitions in self-driven many-particle systems and related non-equilibrium models: A network approach.* J. Statist. Phys., 112, pp.135-153 (2003).

[4] Almgren, F., Taylor, J. E. and Wang, L. *Curvature-driven flows: a variational approach.* SIAM J. Control Optim. 31 387-438, (1993).

[5] Ambrosio, L. and Gigli, N., *A user's guide to optimal transport.* CIME summer school of Cetraro (2009).

[6] Ambrosio L., Gigli N., and Savare G. *Gradient flows in metric spaces and in the space of probability measures.* Lecture in Mathematics ETH Zurich. Birkhuser Verlag, Basel (2005).

[7] Arnold, V. I. and Khesin, B. A. *Topological methods in hydrodynamics.* Springer-Verlarg, New York (1998).

[8] Arnold, V. I. *Mathematical Methods of Classical Mechanics.* Springer-Verlarg, New York (1989).

[9] Balmforth, N. J. and Sassi, R.: *A shocking display of synchrony.* Physica D **143**, 21-55 (2000).

[10] Bakry,D. and Emery, M. *Diffusions hypercontractives.* In "Sem. Probab., XIX," Lecture Notes in Math., Vol. 1123, pp. 177–206, Springer-Verlag, New York Berlin (1985).

[11] Benamou, J.-D and Brenier, Y., *A computational fluid mechanics solution to the Monge-Kantorovich mass transfer problem.* Numer. Math., 84 pp.375-393 (2000).

[12] Brenier, Y. *Décomposition polaire et réarregenment monotone des champs de vecteurs.* C. R. Acad. Sci. Paris Sér. I math. 305, 19, 805-808 (1987).

[13] Brenier, Y. *Polar factorization and monotone rearrangement of vector-valued functions.* Comm. Pure. Appl. Math. 44, **4** 375-417 (1991).

[14] Benedetto, D., Caglioti, E. and Montemagno, U.: *On the complete phase synchronization for the Kuramoto model in the mean-field limit.* To appear in Communications in Mathematical Sciences.

[15] Ben-Jacob, E., Cohen,I., Cziró,k A., Shochet, O. and Vicsek , T. *Novel type of phase transition in a system of self-driven particles.* Phys. Rev. Lett. 75, pp. 1226-1229 (1995).

[16] Bolley F., Cañizo J. A., abd Carrillo J. A. *Mean-field limit for the stochastic Vicsek model.* Appl. Math. Lett. 25, pp. 339-343 (2012).

[17] Bostan, M. and Carrillo, J. A. *Asymptotic fixed-speed reduced dynamics for kinetic equations in swarming* Math. Models Methods Appl. Sci., 23, pp. 2353-2393 (2013).

[18] Buck, J. and Buck, E.: *Biology of synchronous flashing of fireflies.* Nature **211**, 562 (1966).

[19] Carlen, E. A., and Gangbo, W. *Solution of a Model Boltzmann Equation via Steepest Descent in the 2-Wasserstein Metric.* Arch. Ration. Mech. Anal. 172, 1, 21-64 (2004).

[20] Carrillo, J. A., Choi, Y.-P., Ha, S.-Y., Kang, M.-J. and Kim, Y.: *Contractivity of transport distances for the kinetic Kuramoto equation.* J. Stat. Phys. **156**, 395-415 (2014).

[21] Carrillo J., Lisini S., Savare G., and Slepcev D., *Nonlinear mobility continuity equations and generalized displacement convexity.* J. Funct. Anal, 258, no. **4** 1273-1309 (2010).

[22] Chiba, H.: *A proof of the Kuramoto conjecture for a bifurcation structure of the infinite-dimensional Kuramoto model.* Ergodic Theory Dynam. Systems **35**, (2015), 762834 (2015).

[23] Chiba, H.: *Continuous limit of the moments system for the globally coupled phase oscillators.* Discrete Contin. Dyn. Syst. **33** (2013), 18911903 (2013).

[24] Choi, Y.-P., Ha, S.-Y., Jung, S. and Kim, Y.: *Asymptotic formation and orbital stability of phase-locked states for the Kuramoto model.* Physica D, **241**, 735-754 (2012).

[25] Chopra, N. and Spong, M. W.: *On exponential synchronization of Kuramoto oscillators.* IEEE Trans. Automatic Control **54**, 353-357 (2009).

[26] Couzin, I. D., Krause, J., James, R., Ruxton, G. D., and Franks, N. R. *Collective memory and spatial sorting in animal groups* J. Theor. Biol. 218, pp. 1-11 (2002).

[27] Degond P., Dimarco, G., and Mac, T. B. N. *Hydrodynamics of the Kuramoto-Vicsek model of rotating self-propelled particles*, Math. Models Methods Appl. Sci., 24, pp. 277-325 (2014).

[28] Degond, P., Frouvelle A., and Liu, J.-G. *Phase transitions, hysteresis, and hyperbolicity for self-organized alignment dynamics* Arch. Rational Mech. Anal., 216, pp. 63-115 (2015).

[29] Degond, P., Frouvelle, A., and Liu, J.-G. *Macroscopic limits and phase transition in a system of self-propelled particles.* J. Nonlinear Sci., 23, pp. 427-456 (2012).

[30] Degond, P., Frouvelle, A., and Merino-Aceituno, S. *A new flocking model through body attitude coordination.* http://arxiv.org/abs/1605.03509 (2016).

[31] Degond, P. and Motsch, S. *Continuum limit of self-driven particles with orientation interaction.* Math. Models Methods Appl. Sci., 18, pp. 1193-1215 (2008).

[32] Degond, P. and Motsch, S. *Macroscopic limit of self-driven particles with orientation interaction.* C. R. Acad. Sci. Paris, Ser I. 345, pp. 555-560 (2007).

[33] Degond, P. and Yang, T. *Diffusion in a continuum model of self-propelled particles with alignment interaction.* Math. Models Methods Appl. Sci., 20, pp. 1459-1490 (2010).

[34] Denis R. Akhmetov, Mikhail M. Lavrentiev, Jr., and Renato Spigler. *Existence and uniqueness of classical solutions to certain nonlinear integro-differential Fokker-Planck type equations.* Electronic Journal of Diferential Equations, Vol. **2002**(2002) No.24, pp. 1-17.

[35] Dong, J.-G. and Xue, X.: *Synchronization analysis of Kuramoto oscillators.* Commun. Math. Sci. **11**,465-480 (2013).

[36] Dörfler, F. and Bullo, F.: *Synchronization in complex networks of phase oscillators: A survey.* Automatica **50**, 1539- 1564 (2014).

[37] Dörfler, F. and Bullo, F.: *On the critical coupling for Kuramoto oscillators.* SIAM. J. Appl. Dyn. Syst. **10**,1070-1099 (2011).

[38] Fathi, A. *Théorème KAM faible et theorie de Mather sur les systeme lagrangiens.* C. R. Acad. Sci. Paris Sr. I Math. 324, 1043-1046 (1997).

[39] Fathi, A. *Weak KAM Theorem and Lagrangian Dynamics* Cambridge Studies in Advanced Mathematics, (2016).

[40] Fathi, A. and Figalli, A. *Optimal transportation on non-compact manifolds.* Israel J. Math. 175, pp. 1-59 (2010).

[41] Fernandez, B., Gérard-Varet, D. and Giacomin, G.: *Landau damping in the Kuramoto model.* Annales Henri Poincaré **17**, 1793-1823 (2016).

[42] Figalli, A., Gangbo, W., and Turkay Y. *A variational method for a class of parabolic PDEs*, Ann. Sc.Norm. Super. Pisa Cl. Sci. (5)10, 207-252 (2011).

[43] Figalli, A. and Gigli, N. *A new transportation distance between non-negative measures, with appications to gradient flows with Dirichlet boundary Conditions*, J. Math Pures Appl.(9),94 (2),107-130 (2010).

[44] Figalli, A., Kang, M.-J., and Morales, J. *Global well-posedness of spatially homogeneous Kolmogorov-Vicsek model as a gradient flow.* submitted, http://arxiv.org/pdf/1509.02599 (2015).

[45] Figalli, A., and Villani, C. *Optimal Transport and Curvature.* Nonlinear PDE's and applications, 171-217, Lecture Notes in Math., 2028, Springer, Heidelberg (2011).

[46] Frouvelle, A. *A continuum model for alignment of self-propelled particles with anisotropy and density dependent parameters.* Math. Mod. Meth. Appl. Sci. 22, 1250011 (2012).

[47] Frouvelle, A. and Liu, J.-G. *Dynamics in a kinetic model of oriented particles with phase transition.* SIAM J. Math. Anal. 44, 791-826 (2012).

[48] Gamba, I. M., Haack, J. R., and Motsch, S. *Spectral method for a kinetic swarming model.* Journal of Computational Physics, 297, pp. 32-46 (2015).

[49] Gamba, I. M., and Kang, M.-J. *Global weak solutions for Kolmogorov-Vicsek type equations with orientational interaction.* Archive for Rational Mechanics and Analysis, 222(1), 317-342 (2016).

[50] Grégoire, G. and Chaté, H. *Onset of collective and cohesive motion.* Phys. Rev. Lett. 92, 025702 (2004).

[51] Gozlan, N., and Leonard, C. *Transport inequalities. A survey.* http://arxiv. org/abs/1003.3852 (2010).

[52] Ha, S.-Y., Jeong, E., and Kang, M.-J. *Emergent behaviour of a generalized Viscek-type flocking model.* Nonlinearity, 23, pp. 3139-3156 (2010).

[53] Ha, S.-Y., Kim, H. K., Morales, J., and Park, J. *Emergence of phase concentration for the Kuramoto-Sakaguchi equation. arXiv:* 1610.01703 [math.AP], (2016.)

[54] Ha, S.-Y., Ko, D., Park, J. and Zhang, X.: *Collective synchronization of classical and quantum oscillators.* To appear in EMS Surv. Math. Sci.

[55] Ha, S.-Y., Kim, H. K., and Park, J.: *Remarks on the complete synchronization of Kuramoto oscillators.* Nonlinearity. **28**, 1441-1462 (2015).

[56] Ha, S.-Y., Kim, H. K. and Ryoo, S. W.: *On the finiteness of collisions and phase-locked states for the Kuramoto model.* Submitted.

[57] Ha, S.-Y., Kim, H. K. and Ryoo, S. W.: *Emergence of phase-locked states for the Kuramoto model in a large coupling regime.* Commun. Math. Sci. **14**, 1073-1091 (2014).

[58] Ha, S.-Y., Ko, D., Park, J. and Zhang, X.: *Collective synchronization of classical and quantum oscillators.* Submitted.

[59] Ha, S.-Y., Li, Z. and Xue, X.: *Formation of phase-locked states in a population of locally interacting Kuramoto oscillators.* J. Differential Equations. **255**, 3053-3070 (2013).

[60] Ha, S.-Y., Noh, S. E. and Park, J.: *Practical synchronization of generalized Kuramoto systems with an intrinsic dynamics.* Netw. Heterog. Media **10**, 787807 (2015).

[61] Ha, S.-Y. and Xiao, Q.: *Nonlinear instability of the incoherent state for the Kuramoto-Sakaguchi-Fokker-Plank equation.* J. Stat. Phys. **160**, 477-496 (2015).

[62] Ha, S.-Y. and Xiao, Q.: *Remarks on the nonlinear stability of the Kuramoto-Sakaguchi equation.* J. Differential Equations **259**, 2430-2457 (2015).

[63] Holley, R. and Stroock, D. *logarithmic Sobolev inequalities and stochastic Ising models.* J. Statist. Phys. 46, 1159-‘1194 (1987).

[64] A. D. Ioffe, *On lower semicontinuity of integrand functionals.* I. SIAM J. Control Optim., 15, 521-538 (1977).

[65] Jadbabaie, A., Motee, N. and Barahona M.: *On the stability of the Kuramoto model of coupled nonlinear oscillators.* Proceedings of the American Control Conference, 4296-4301 (2004).

[66] Jordan R., Kinderlehrer, D., and Otto, F.: *The variational formulato of the Fokker-Planck equations.* SIAM J. Math. Anal., 29 (1998) no. 1, 1-17.

[67] M. Kang and J. Morales. *Dynamics of a spatially homogeneous Vicsek model for oriented particles on the plane. arXiv:* 1608.00185 [math.AP], (2016).

[68] Kennard, E. H.: Kinetic theory of gases. McGraw-Hill Book Company, New York and London 1938.

[69] Kuramoto, Y.: *Chemical Oscillations, waves and turbulence.* Springer-Verlag, Berlin. 1984.

[70] Kuramoto, Y.: *International symposium on mathematical problems in mathematical physics.* Lecture notes in theoretical physics. **30**, 420 (1975).

[71] Lancellotti, C.: *On the Vlasov limit for systems of nonlinearly coupled oscillators without noise.* Transport theory and statistical physics. **34**, 523-535 (2005).

[72] Liero, M., Mielke, A., and Savare, G. *Optimal Entropy-Transport problems and a new Hellinger-Kantorovich distance between positive measures. arXiv:* 1508.07941 [math.OC], 2015.

[73] Liu, T.-P. and Yu, S.-H.: *Boltzmann equation: micro-macro decompositions and positivity of shock profiles.* Comm. Math. Phys. **246** (2004), 133179 (2004).

[74] Liu, T.-P. and Zeng, Y.: *Time-asymptotic behavior of wave propagation around a viscous shock profile.* Comm. Math. Phys. **290**, 2382 (2009).

[75] Mather, J. N. *Minimal measures.* Comment. Math. Helv. 64,3 (1989), 275-394.

[76] Mather, J. N. *Action minimizing invariant measures for positive definite Lagrangian systems.* Math. Z. 207,2, 169=207 (1991).

[77] McCann, R. J. *Polar factorization of maps on Riemannian manifolds.* Geom. Funct. Anal., 11, pp. 589-608 (2001).

[78] Milnor, J. *Morse Theory.* Annals of Mathematics Sudies Number 51 (1963).

[79] Morales, J. *A new family of transportation costs with applications to reaction-diffusion and parabolic equations with boundary conditions. arXiv:* 1604.00551 [math.AP], (2016).

[80] Monsaingeon, L., Kondratyev, S., and Vorotnikov, D. *A new optimal transport distance on the space of finite Radon measures. arXiv:* 1505.07746 [math.AP], (2015).

[81] Neunzert, H.: *An introduction to the nonlinear Boltzmann-Vlasov equation. In kinetic theories and the Boltzmann equation.* Lecture Notes in Mathematics 1048, Springer, Berlin, Heidelberg.

[82] Otto, F. *Evolution of microstructure in unstable porous media: a relaxational approach.* Comm. Pure Appl. Math., 52 no 7.873-915, (1999).

[83] Otto, F.: *The geometry of dissipative evolution equations: the porous medium equation.* Comm. PDE. **26**(1-2), 101-174 (2001).

[84] Otto F., and Villani C. *Generalization of an inequality by Talagrand and links with the Logarithmic Sobolev inequality* J. Funct. Anal. 173, 361-400 (2000).

[85] Otto, F. and Tzavaras, A. *Continuity of velocity gradients in suspensions of rod-like molecules.* Commun. Math. Phys. 277, pp. 729-758 (2008).

[86] Piccoli B. and Rossi F. *A generalized Benamou-Brenier formula for mass-varying densities.* arXiv:1304.7014v1[math.Ap] (2013).

[87] Pikovsky, A., Rosenblum, M. and Kurths, J.: *Synchrnization: A universal concept in nonlinear sciences.* Cambridge University Press, Cambridge, 2001.

[88] Rifford, L. *Regularity of weak KAM solutions the Mañe's Conjecture* Séminaire Laurant Schawrtz-EDP et applications, Talk no. 27, p. 1-22, (2011-2012).

[89] Ronghua, P. *Darcy's law in one-dimensional Isentropic Porous medium flow.* Series in Contemporary Applied Mathematics, CAM 17, Vol 1.,238–250, (2012).

[90] Rothaus, O. *Diffusion on compact riemannian manifolds and logarithmic sobolev inequalities.* J. Funct. Anal. 42, 102-109 (1981).

[91] Simon J. *Compact Sets in the Space* $L^p(0,T;B)$. Ann. Mat. Pura Appl., 196 (1987).

[92] Van Hemmen, J. L. and Wreszinski, W. F.: *Lyapunov function for the Kuramoto model of nonlinearly coupled oscillators.* J. Stat. Phys. **72**, 145-166 (1993).

[93] Villani, C.: *Topics in optimal transportation.* Graduate Studies in Math. **58**, AMS 2003.

[94] Villani,C. *Optimal transport, old and new.* Grundlehren des mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], Vol 338, Springer-Verlag, Berlin-New York (2009).

[95] Winfree, A. T.: *Biological rhythms and the behavior of populations of coupled oscillators.* J. Theor. Biol. **16**, 15-42 (1967).

[96] Yoon, S. Sorbaro Sindaci, M.Goltsec, A. V. and Mendes, J. F F.: *Critical behavior of the relaxation rate, the susceptibility, and pair correlation function in the Kuramoto model on scale-free networks.* Phys. Rev. E **91**, 032814 (2015).